

## **Tunnisteoppaan laatiminen yrityksen dokumenttienhallinnan tueksi**

Laura Kankaanpää

<b>Tekijä tai tekijät</b> Laura Kankaanpää	
<b>Koulutusohjelma</b> Johdon assistenttityön ja kielten koulutusohjelma	
<b>Raportin nimi</b> Tunnisteoppaan laatiminen yrityksendokumenttienhallinnan tueksi	<b>Sivu- ja liitesivumäärä</b> 42 + 2
<b>Opettajat tai ohjaajat</b> Heta-Liisa Malkavaara	
<p>Työntekijät käyttävät nykypäivänä jopa kaksikymmentä prosenttia työajastaan tiedon etsimiseen eikä tietoa löydetä monesti ollenkaan, vaikka se olisi saatavilla. Tiedon tekeminen helpommin saatavaksi vähentää sen etsimiseen kuluvaa aikaa ja jokaisen organisaation olisi siksi järkevää panostaa tiedonhallinnan kehittämiseen.</p> <p>Tämä toiminnallinen opinnäytetyö keskittyy sähköisen dokumenttienhallinnan erilaisiin menetelmiin ja pohjautuu Bluefors Oy:n antamaan toimeksiantoon. Kohdeyrityksen toimeksianton tavoitteena oli määritellä yhtenäiset tavat metatietoihin pohjautuvan dokumenttienhallintatyökalun tunnisteiden luomiseen. Määrittelyprosessin pohjalta Bluefors Oy:lle luotiin kirjallinen tunnisteopas, jonka avulla työntekijät voivat olla osana yrityksen onnistunutta dokumenttien hallintaa.</p> <p>Raportin tietoperusta antaa lukijalle yleiskäsityksen metatiedoista sekä tunnisteiden käytöstä tiedonhallinnan työkaluna. Lukija voi oivaltaa näiden sivujen pohjalta metatietojen yhteyden jokapäiväiseen elämään työssä ja vapaa-ajalla.</p> <p>Tietoperustan lisäksi opinnäytetyössä on esitelty kohdeyritykselle toteutettu tunnisteiden kielen sekä rakenteen määrittelyprosessi, joka sai alkunsa toukokuussa 2019 ja tuli päätökseen lokakuussa 2019. Prosessin aikana syntynyt tunnisteopas on esitelty prosessia seuraavassa luvussa. Tunnisteopas on luottamuksellinen, jonka vuoksi ainoastaan oppaan sisällysluettelo sekä oppaan lopussa sijaitseva pikaopas löytyvät tästä raportista.</p>	
<b>Asiasanat</b> Metatieto, tunniste, dokumenttienhallinta, tiedon hakutavat, opas	

# Sisällys

1	Johdanto.....	1
1.1	Toimeksiannon esittely ja rajaus.....	2
1.2	Bluefors Oy:n esittely .....	3
2	Tieto ja sen hallinta .....	5
2.1	Tieto ja metatieto .....	5
2.1.1	Metatiedon eri lajit.....	6
2.1.2	Metatiedon neljäs laji .....	9
2.2	Tiedonhallinta.....	11
3	Prosessin elinkaari ja taustamääritelmät.....	14
3.1	Oppaan laatiminen .....	16
3.2	Hakutavan ymmärtäminen .....	17
3.2.1	Erilaiset hakutavat.....	18
3.2.2	Hakutavan valinta .....	21
4	Tunnisteoppaan esittely .....	22
4.1	Tunnisteiden kieli .....	23
4.1.1	Kielen valinta.....	24
4.1.2	Ei-englanninkielisissä dokumenteissa käytettävä tunnistekieli .....	26
4.2	Merkistöstä yhtenäisyyttä tunnisteisiin .....	27
4.2.1	Merkistön tutkiminen .....	28
4.2.2	Merkistön määrittelemineen.....	29
4.3	Tunnisteiden muodostaminen.....	31
4.4	Tunnistekategoriat .....	34
5	Kehittäminen ja arviointi .....	36
5.1	Prosessin kehityskohtia .....	36
5.2	Projektinhallinta.....	37
5.3	Määrittelyprosessin ja oppaan arviointi .....	37
	Lähteet .....	39
	Liitteet .....	43
	Liite 1. Guidelines for Tagging -oppaan sisällysluettelo .....	43
	Liite 2. Visuaalinen pikaopas tunnisteiden luomiseen .....	44

## Kuvat

Kuva 1: Tietoperusta-luvun rakenne .....	2
Kuva 2: Bluefors Oy:n monikansalaisuus prosentteina (Hanna Puura 25.6.2019) .....	4
Kuva 3: Luomispäivä (Date created) ja muokkauspäivä (Date modified) ovat metatietoja. .6	
Kuva 4: Authors kenttään ilmestyy automaattisesti tiedostonluoja, kun tiedosto juodaan Microsoft Office -ohjelmaa käyttäen.....	7
Kuva 5: Sisällöllinen metatieto tag-sarakkeessa.....	8
Kuva 6: Instagram-julkaisun hashtag-aihetunnisteet ovat tunnisteita (Kankaanpää 2017) .9	
Kuva 7: Synonyymiringin toimintaperiaate.....	11
Kuva 8: Kaarion ja Peltolan kuvaus tiedon elinkaaren päävaiheista.....	13
Kuva 9: Prosessin ajanjakso ennen ja jälkeen toimeksiannon .....	14
Kuva 10: Oppaan laatimisprosessin vaiheet.....	16
Kuva 11: Käyttäjä näkee dokumenttirakenteet helposti kansirakenteesta .....	18
Kuva 12: Esimerkki kansiohierarkian kasvamisesta liian pitkäksi.....	19
Kuva 13: Tiedoston nimi kertoo tietoa dokumentista ja sen sisällöstä.....	20
Kuva 14: Tunnisteoppaan aiherakenne .....	22
Kuva 15: Yksikkö- ja monikkomuotojen käyttö tunnisteissa.....	32

## Taulukot

Taulukko 1: Sanaston eroavaisuuksia amerikanenglannin ja brittienglannin välillä (Beare 2019).....	24
Taulukko 2: Kirjoitusasun eroja amerikanenglannin ja brittienglannin välillä (British Council 2019).....	24
Taulukko 3: Esimerkkejä kielitunnisteista .....	27
Taulukko 4: Tunnisteisiin hyväksytyt suuraakkoset .....	29
Taulukko 5: Tunnisteisiin hyväksytyt pienaakkoset .....	30
Taulukko 6: Tunnisteisiin hyväksytyt numerot .....	30
Taulukko 7: Ainut tunnisteisiin hyväksytty erikoismerkki.....	30
Taulukko 8: Esimerkkejä erikoisaakkosten korvaamiseen tunnisteissa.....	30
Taulukko 9: kahdeksan esimerkkiä yhden tunnisteiden kirjoitusasuvariaatioista .....	31
Taulukko 10: Vertailu yksittäisten konseptien kirjoitustavoista suomeksi ja englanniksi ...	33
Taulukko 11: Tunnisteiden muodostaminen alaviivaa käyttäen.....	34
Taulukko 12: Tunnistekategoriat ja niiden selitykset.....	35

# 1 Johdanto

McKinsey Global Instituten (MGI) tekemän tutkimuksen mukaan vuorovaikutustyöntekijät käyttävät keskimäärin lähes kaksikymmentä prosenttia työajastaan etsien tietoa tai jäljitellen kollegoita, jotka voisivat olla avuksi tietyssä työtehtävissä (Chui & ym. 2012). Vuorovaikutustyöntekijöillä tarkoitetaan koulutettuja ammattilaisia, kuten johtajia, myyjiä ja asiakaspalvelijoita, jotka käyttävät suuren osan työajastaan vuorovaikutuksessa ihmisten kanssa (Trends Magazine 2013). Kärjistettynä MGI:n tekemä tutkimus tarkoittaa viiden henkilön myyntitiimissä kokonaisen henkilön työajan kulumista pelkkään tiedon etsimiseen ilman muun työpanoksen antamista työyhteisölle (Di Noi 2018).

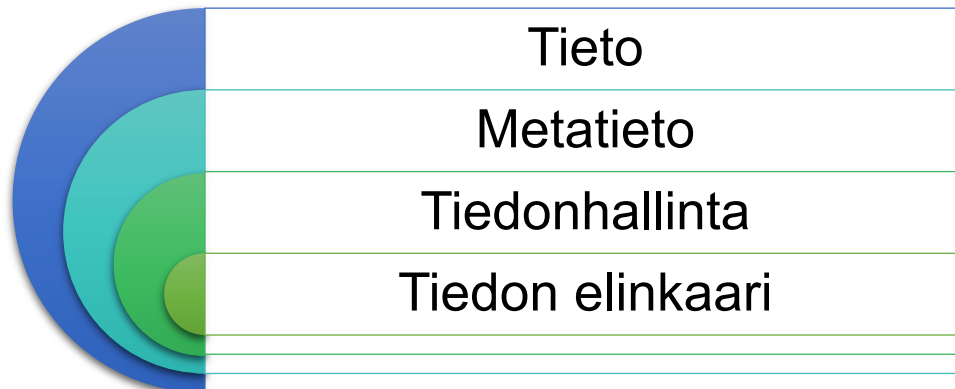
Tämän opinnäytetyön tavoitteena oli tutustua erilaisiin dokumenttienhallinnan tapoihin, joiden avulla pystyttäisiin vähentämään Bluefors Oy:ssä tiedon etsimiseen kuluva aikaa. Lisäksi opinnäytetyön prosessin aikana luotiin opas kohdeyrityksen uuden dokumenttienhallintatyökalun tueksi, joka tulee helpottamaan dokumenttienhallintatyökalun käyttöä sekä Bluefors Oy:n työntekijöiden työtä.

Raportissa esitellyt dokumenttien hallinnan menetelmät keskittyvät sähköisten tiedostojen sekä dokumenttien hallintaan, jotka ovat toteutettavissa ilman ulkopuolisten toimijoiden tarjoamia palveluita. Tässä opinnäytetyössä käsitellään erityisesti metatietoihin perustuvaa dokumenttienhallintaa, joka valittiin Bluefors Oy:lle sopivaksi dokumenttienhallinnan menetelmäksi toimeksiannon yhteydessä. Tarkempi toimeksiannon kuvaus sekä kohdeyrityksen esittely löytyvät tästä luvusta raportin rakenteen sekä keskeisten sanojen esittelyn jälkeen.

Opinnäytetyöraportin aikana esitellään tietoperustan lisäksi prosessia, jonka aikana hankin tietoa erilaisista dokumenttienhallinnan ratkaisuksista, määrittelin kohdeyritykselle räätälöidyt käytännöt valitun ratkaisutavan tueksi sekä laadin oppaan näiden määrittelyjen selkeyttämiseksi kohdeyrityksen työntekijöille. Näistä työvaiheista koostuvaa prosessia kutsutaan tässä opinnäytetyössä määrittelyprosessiksi sekä prosessiksi. Prosessin lisäksi esittelen määrittelyprosessin aikana syntyneen tunnisteoppaan, joka sai kohdeyrityksessä nimekseen Guidelines for Tagging. Guidelines for Tagging -oppaasta puhutaan oppaan nimen lisäksi sanoilla opas, tunnisteopas sekä prosessin tuotos. Määrittelyprosessin kuvaus sekä aikataulu löytyvät luvusta kolme, jota seuraa tunnisteoppaan esittely luvussa neljä.

Prosessin ja sen tuotoksen esittelyä edeltää luvun kaksi tietoperusta, joka käsittelee tietoa ja sen hallintaa. Tietoperusta alkaa tiedon kuvaamisella ja esittelee käsitteitä metatieto,

tiedonhallinta sekä tiedon elinkaari (kuva 1). Metatiedon yhteydessä käsitellään metatiedon eri lajeja sekä metatiedon neljännessä lajissa, tunnisteissa, käytettävän kielen merkitystä, joka oli määrittelyprosessin keskeisin teema.



Kuva 1: Tietoperusta-luvun rakenne

Tietoperustan, määrittelyprosessin kuvaamisen ja tunnisteoppaan esittelyn lisäksi raportti sisältää luvun, jossa arvioin prosessia ja Guidelines for Tagging -oppaan onnistumista. Samassa luvussa kerron myös prosessiin liittyvistä kehityskohdista sekä -ideoista, joiden avulla voidaan mahdollisesti tukea esimerkiksi teknisillä ratkaisuilla määriteltujen toimintatapojen toteutumista. Nämä kehitysideat voivat olla hyödyllisiä samankaltaisten toimeksiantojen toteuttamisessa sekä Bluefors Oy:n dokumenttienhallinnan kehittämisessä tulevaisuudessa.

### 1.1 Toimeksiannon esittely ja raja

Bluefors Oy:n antama toimeksianto oli määritellä yrityksen käyttöön tulevassa dokumenttienhallintatyökalussa käytettävien tunnisteiden kieltä ja rakennetta koskevat piirteet. Määrittelyjen lisäksi toimeksiantoon sisältyi yrityksen työntekijöiden käyttöön tulevan tunnisteoppaan kirjoittaminen, jonka avulla työkalun käyttäjät saataisiin noudattamaan yhdenmukaisesti prosessin aikana tehtyjä määritelmiä.

Kohdeyritys antoi toimeksiannon toukokuussa 2019. Toimeksianto tuli ajankohtaiseksi keväällä 2019, kun noin kymmenelle Bluefors Oy:n vuorovaikutustyöntekijälle järjestettävässä kevyessä haastattelussa selvisi, että olemassa olevien tiedon löytäminen on työntekijöille haastavaa ja aikaa vievää. Olemassa olevalla tiedolla tarkoitetaan tässä tapauksessa informaatiota, joka on tallennettu kirjallisena yrityksen intranettiin tai tiedostopalvelimelle tallennettuihin tiedostoihin. Haastattelussa kysyttiin myös haastateltavien mielipidettä työkalusta, jonka avulla tietoa voisi hakea helposti esimerkiksi hakusanoja käyttämällä. Suurin osa haastatteluun osallistuneista piti tällaista työkalua hyvänä ideana, joka

johti dokumenttienhallintatyökalun ideointiin sekä tässä opinnäytetyössä esiteltävään prosessiin ja tunnisteoppaan kirjoittamiseen. Prosessin taustat sekä aikataulu on kuvattu tarkemmin luvussa kolme ”Määrittelyprosessi”.

Vuorovaikutustyöntekijöille kohdistetun haastattelukierroksen ja MGI:n tutkimustulosten lisäksi toimeksianto koettiin tärkeäksi Bluefors Oy:ssä vuoden 2019 aikana implementoidun Lean 5S-menetelmän vuoksi. Lean 5S-menetelmän avulla on tarkoitus luoda työpaikkojen organisointiin sekä työmenetelmiin standardeja, joilla nopeutetaan ja helpotetaan työntekoa (Arter Oy 2019; Lean Lion Oy 2019). Nämä tavoitteet ovat samat tässä raportissa kuvailtavan prosessin ja tunnisteoppaan kohdalla. Uuden dokumenttienhallintatyökalun ja sen tueksi kirjoitetun oppaan avulla pystyttäisiin implementoida Lean 5S-menetelmät kesän 2019 aikana standardoitujen fyysisen työympäristön lisäksi myös Bluefors Oy:n sähköiseen työskentely-ympäristöön.

Kohdeyrityksen uuden dokumenttienhallintatyökalun valinta ja tekninen toteutus eivät olleet osa saamaani toimeksiantoa. Tutustuin kuitenkin prosessin aikana erilaisiin dokumenttienhallinnan tapoihin ymmärtääkseni aiheita paremmin sekä varmistuakseni siitä, että ennalta määritelty metodi oli paras mahdollinen kohdeyrityksen tarpeisiin. Vaihtoehtoisia dokumenttienhallinnan tapoja on käsitelty prosessikuvauksen yhteydessä luvussa kolme. Raportin viimeisessä luvussa esitetään prosessin aikana tekemiäni pohdintoja teknisistä tavoista, joiden avulla tunnisteoppaaseen määriteltyjä kielellisiä piirteitä voitaisiin tukea. Näitä kahta poikkeusta lukuun ottamatta keskityin prosessin aikana määrittelemään dokumenttienhallintatyökalussa käytettävien tunnistesanojen kielellisiä sekä rakenteellisia ominaisuuksia, kuten niissä käytettävä kieli ja tunnisteiden muodostamiseen liittyvät säännöt.

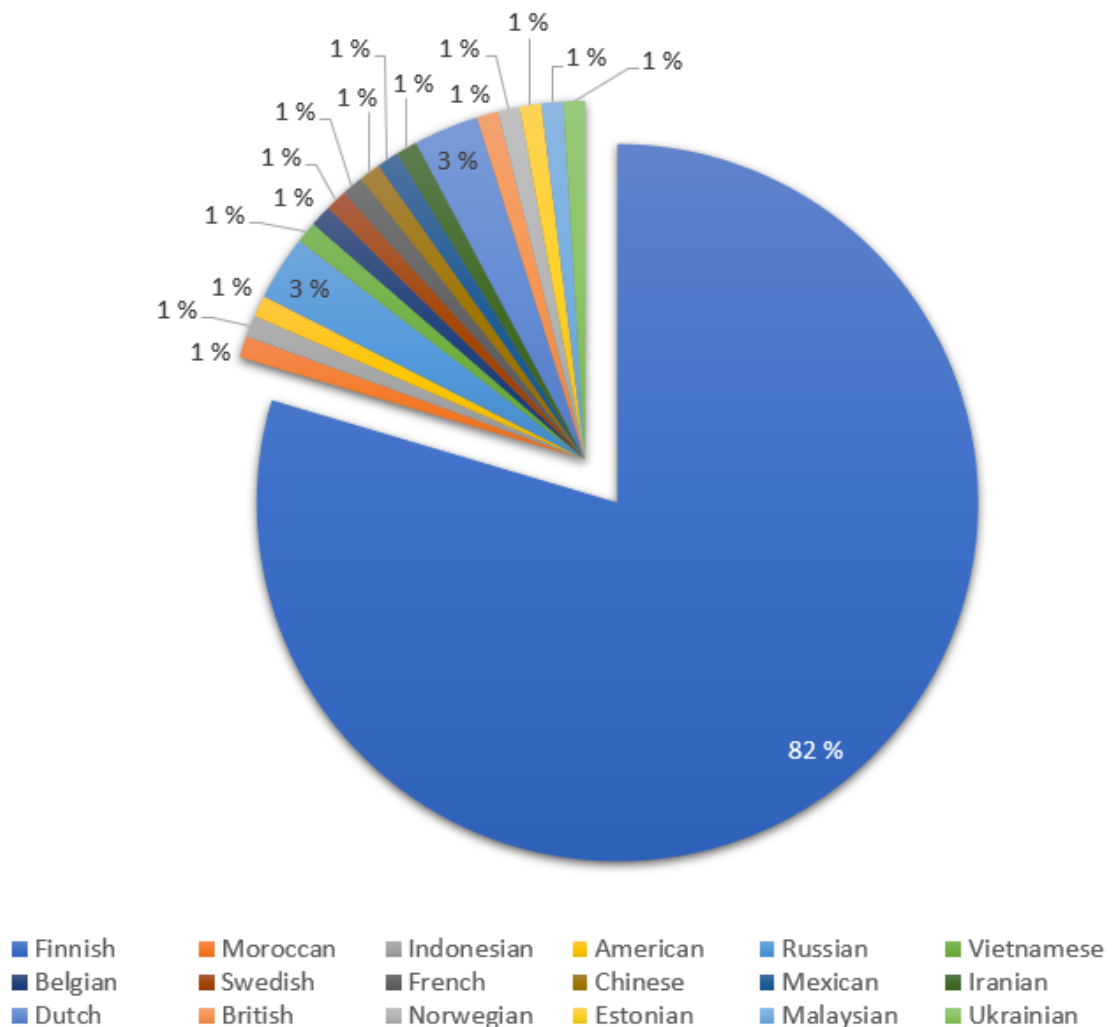
Tämän opinnäytetyön produktio on toteutettu ainoastaan kohdeyrityksen sisäiseen käyttöön ja tunnisteopas sisältää sellaista tietoa, joka on luokiteltu salassa pidettäväksi. Tämän vuoksi ainoastaan Guidelines for Tagging -tunnisteoppaan sisällysluettelo sekä oppaan sisältämä visuaalinen pikaopas ovat nähtävillä raportin lopussa (liite 1 & liite 2).

## **1.2 Bluefors Oy:n esittely**

Opinnäytetyön produktion kohdeyritys on suomalainen Bluefors Oy. Bluefors Oy on noin 160 työntekijää työllistävä teollisuusalan yritys, joka suunnittelee ja rakentaa kryostaatteja teknologian tutkimuksen sekä kehittämisen huippuyrityksille ympäri maailmaa. Hollantilaisen Rob Blaauwgersin sekä Pieter Vorselmanin vuonna 2008 perustama yritys on saanut alkunsa Otaniemessä sijaitsevasta Aalto-yliopiston kylmälaboratoriosta, josta lähes kaikki yrityksen toiminnot ovat muuttaneet Helsingin Pitäjänmäkeen. Pitäjänmäessä toimivan

pääkonttorin ja tuotantolaitoksen lisäksi Bluefors Oy:llä on vuonna 2018 avattu toimisto Münchenissä Saksassa, tytäryhtiö Bluefors Inc. USA:ssa sekä toimintaa Hollannissa (Bluefors 2019a; Bluefors 2019b). (Bluefors 2019c.)

Kansainvälisen yritystoiminnan ansiosta Suomen tasavallan presidentti Sauli Niinistö myönsi Bluefors Oy:lle kansainvälistymispalkinnon vuonna 2016 (Yle 2016). Kansainvälisen liiketoiminnan lisäksi yritys on myös työntekijöidensä puolesta hyvin monikulttuurinen. Kesäkuussa 2019 Bluefors Oy:ssä työskenteli yrityksen henkilöstöosaston mukaan ihmisiä yli 18 eri maasta (Puura 25.6.2019). Kansalaisuuksien prosentuaalisen jaon pystyy näkemään alla olevasta kuvasta (kuva 2). Osalla yrityksen työntekijöistä on monikansalaisuuksia. Tämän vuoksi alla olevassa kuvassa prosenttien summa on yli sata.



Kuva 2: Bluefors Oy:n monikansalaisuus prosentteina (Hanna Puura 25.6.2019)

Tässä opinnäytetyössä kuvatus prosessin tuotos tulee käyttöön yli 50 Bluefors Oy:n työntekijälle, jotka ovat vastuussa yrityksen dokumenttien luomisesta sekä päivittämisestä



(Bluefors 2019d). Näihin henkilöihin kuuluvat muun muassa tuotannon esimiehet ja asiantuntijat, jotka vastaavat työohjeiden luomisesta, sekä kaikki toimiston puolella työskentelevät henkilöt. Näiden yli 50 päätetyöntekijän ja esimiehen lisäksi, prosessin tuotos tulee auttamaan aivan jokaista kohdeyrityksen työntekijää. Uusi dokumenttienhallintatyökalu mahdollistaa dokumenttien hakemisen kaikilta yrityksen sisäisiltä tietokoneilta, joille jokaisella Bluefors Oy:ssä työskentelevällä henkilöllä on pääsy. Haettavat dokumentit eivät rajoitu ainoastaan aiemmin mainittuihin työohjeisiin, vaan niitä ovat myös kaikki yrityksen sisäisellä tietopalvelimella sijaitsevat dokumentit, kuten sairauspoissaolo-ohjeet sekä dokumenttipohjat (Bluefors 2019e). Kyseessä on näin ollen kaikkien yrityksen työntekijöiden työtä helpottava opas.






## **2 Tieto ja sen hallinta**

### **2.1 Tieto ja metatieto**

Sanalle tieto on annettu erilaisia määritelmiä kautta aikojen. Filosofit ovat määritelleet tiedon esimerkiksi hyvin perusteltuna tosi uskomuksena sekä väittämänä, joka on totta. Suomenkielessä sana tieto on kuitenkin hyvin monimerkityksellinen, eikä se aina sisällä filosofisten määrittelyjen näkökulmaa totuudenmukaisuudesta. Kimmo Kaario ja Tuomo Peltola ovat avanneet tiedon merkitystä jakamalla tiedon käsitteisiin data, informaation, tietämys sekä tieto. Suomenkielessä nämä kaikki sanat voidaan määrittää sanalla tieto, vaikka ne ovat käsitteinä hieman erilaisia. (Kaario & Peltola 2008, 6)

Tässä opinnäytetyössä tieto käsitetään informaationa. Informaatio on Kaarion ja Peltolan määritelmän mukaan viesti, jolle vastaanottaja on antanut jonkin merkityksen (Kaario & Peltola 2008, 6). Informaatiota voi olla esimerkiksi tekstiä tai bittijono, jonka ymmärtäminen antaa tietoa jostakin asiasta (Tieteen termipankki 2016a). Raportissa kuvaillussa prosessissa tiedolla käsitetään tekstin lisäksi myös tiedostoja, jotka voivat olla digitaalisesti tallennettua kuvaa, videota sekä ääntä.

Tieto tiedosta on tiedon metamuoto ja sitä kutsutaan metatiedoksi (Dalkir 2011, 23; Smith 2007). Tieto on metatietoa silloin, kun se kuvailee esimerkiksi tiedon kontekstia, sisältöä, rakennetta ja sen käsittelyä koko sen elinkaaren aikana (Tieteen termipankki 2016b). Sähköisessä tiedonhallinnassa metatieto voi olla yksinkertaisimmillaan tietokoneella luodun tiedoston luontipäivä, jonka pystyy näkemään useimmiten tietokoneen kansiorakenteessa (kuva 3). Luontipäivä kertoo, milloin tiedosto on luotu, eli se kuvaa kyseisen tiedoston luomistapahtumaa ja tiedon elinkaaren alkua. Myös luoja-tieto on metatietoa, kuten tiedoston viimeisin muokkauspäivä ja tiedostotyyppi. (Kaario & Peltola 2008, 25-26.)

Name	Date modified	Type	Date created
 [redacted]	31.05.2019 8.00	File folder	31.05.2019 8.00
 [redacted]	31.05.2019 8.00	File folder	31.05.2019 8.00
 [redacted]	20.08.2019 8.14	File folder	31.05.2019 8.00
 [redacted]	05.01.2018 11.10	Microsoft Word D...	31.05.2019 8.00
 [redacted]	12.03.2018 17.17	KDBX File	31.05.2019 8.00















Kuva 3: Luomispäivä (Date created) ja muokkauspäivä (Date modified) ovat metatietoja.

Metatietoa pidetään nykypäivänä tiedonhallinnan kivijalkana, koska se tarjoaa tiedon jäsentelyyn työkalun nykypäivän informaatioähkyisessä ja lisääntyneen tiedon monimutkais-  
tamassa työympäristössä (Dalkir 2011, 23; Serola 2010, 34). Seuraavissa alaluvuissa on kuvailtu metatietoa vielä tarkemmin sekä esitelty metatiedon eri lajeja.

### 2.1.1 Metatiedon eri lajit

Kuvassa 3 nähtävät tiedoston luonti- ja muokkauspäivä sekä tiedoston tyyppi (type), ovat kaikki niin sanottuja hallinnollisia metatietoja. Hallinnollinen metatieto on informaatiota, joka kuvaa digitaalisen tiedon teknisiä ominaisuuksia sekä tiedoston hallintaan liittyviä tietoja, kuten arkistointitietoja (MerlinOne 2019). Näiden metatietojen tärkein tarkoitus on ni-  
mensä mukaan auttaa käyttäjiä ja organisaatioita hallitsemaan tietoa (Smith 2007).

Iso osa hallinnollisista metatiedoista syntyy automaattisesti tiedostoihin, kun tietoa taltioi-  
daan sähköisiksi dokumenteiksi (Kaario & Peltola 2008, 26.). Toiset tietokoneohjelmat liit-  
tävät hallinnollista metatietoa dokumentteihin automaattisesti vain vähän, mutta jotkin oh-  
jelmat voivat kerätä metatietoa dokumenttiin enemmänkin. Tällaisia enemmän tietoa ke-  
rääviä ohjelmia ovat esimerkiksi kaikki Microsoft Office-ohjelmat, kuten Word, Power-Point  
ja Excel (Microsoft 2019). Nämä ohjelmat osaavat liittää tiedostoon luonti- ja muokkaus-  
päivän sekä tiedostotyyppin lisäksi muun muassa, kuka dokumentin on luonut sekä monta  
sanaa ja merkkiä dokumentissa on (Microsoft 2019) (kuva 4).

Name	Date modified	Type	Date created	Authors
 [redacted]	31.05.2019 8.00	File folder	31.05.2019 8.00	
 [redacted]	31.05.2019 8.00	File folder	31.05.2019 8.00	
 [redacted]	20.08.2019 8.14	File folder	31.05.2019 8.00	
 [redacted]	05.01.2018 11.10	Microsoft Word D...	31.05.2019 8.00	[redacted]
 [redacted]	12.03.2018 17.17	KDBX File	31.05.2019 8.00	
 [redacted]	25.05.2018 9.08	Microsoft Word D...	31.05.2019 8.00	Laura Kankaanpää
 [redacted]	15.08.2018 7.39	Microsoft Word D...	31.05.2019 8.00	Laura Kankaanpää
 [redacted]	15.08.2018 17.31	Microsoft Word D...	31.05.2019 8.00	Laura Kankaanpää
 [redacted]	26.04.2019 10.20	Microsoft Word D...	31.05.2019 8.00	Laura Kankaanpää
 [redacted]	26.04.2019 10.39	Foxit Reader PDF ...	31.05.2019 8.00	
 [redacted]	19.08.2019 13.22	Microsoft Word D...	19.08.2019 13....	Laura Kankaanpää
 [redacted]	19.08.2019 13.22	Microsoft PowerP...	19.08.2019 8.49	Laura Kankaanpää
 [redacted]	16.09.2019 14.28	Microsoft Excel W...	16.09.2019 14....	Laura Kankaanpää
 [redacted]	16.09.2019 15.17	Microsoft Excel W...	16.09.2019 15....	Laura Kankaanpää

Kuva 4: Authors kenttään ilmestyy automaattisesti tiedostonluoja, kun tiedosto juodaan Microsoft Office -ohjelmaa käyttäen.

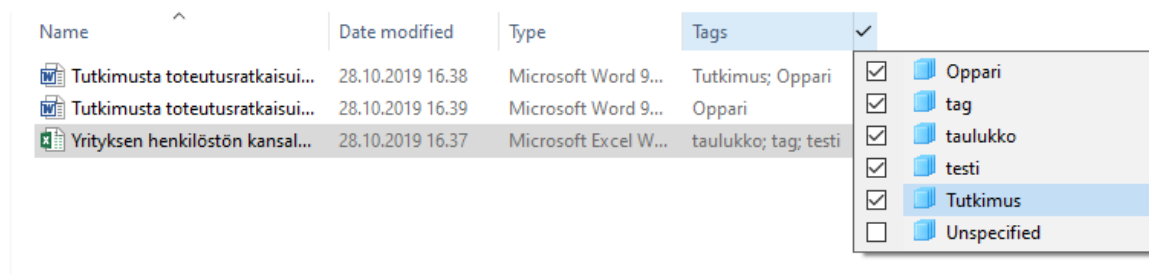
Vaikka useasti fyysisen tiedon metatieto näkyy ensimmäisen kerran tietokoneen ruudulla, se voi muodostua jo ennen tiedon taltiointia tietokoneelle. Esimerkiksi digitaalisten still-kuva- ja videotiedostojen kohdalla hallinnollista metatietoa syntyy jo kuvausvaiheessa, kun tiedosto luodaan suoraan digitaaliseen muotoon. Siinä vaiheessa, kun kuva tallentuu tiedostoksi, siihen syntyy automaattisesti sen luomiseen käytetyn laitteiden tiedot sekä kuvan resoluutio, jotka ovat tietoa kuvan teknisistä ominaisuuksista. (MerlinOne 2019; Digi-kuva 2018.)

Hallinnollista metatietoa on teknisten ominaisuuksien sekä hallintatietojen lisäksi myös dokumentin oikeuksiin liittyvät tiedot, kuten käyttöäioikeudet. Käyttäjäioikeuksilla ei tarkoiteta tässä kohtaa henkilöitä, joille tieto on kohdennettu, vaan henkilöitä, joilla on oikeus kyseisen tiedon näkemiseen sekä tapoja, joilla tiedon saa näkyville. Kaiken kaikkiaan hallinnollisen metatiedon voi jakaa kolmeen eri kategoriaan: tekninen tieto, säilytystieto sekä oikeuksia koskeva tieto. (Dalkir 2011, 210; MerlinOne 2019.)

Hallinnollisten metatietojen lisäksi olemassa on rakenteellista sekä sisällöllistä metatietoa. Rakenteellinen metatieto kuvaa tiedon fyysistä sekä loogista rakennetta ja sen päätarkoitus on luoda suhteita yksittäisten tietojen välillä (Smith 2007). Rakenteellinen metatieto kuvaa esimerkiksi, miten digilehden yksittäiset artikkelit kuuluvat tiettyyn julkaisuun (Lager 2016). Rakenteellinen metatieto kuvaa myös logiikkaa, jolla digitaalisen kirjan sivut osaa- vat mennä oikeaan sivujärjestykseen, vaikka ne olisivat itsenäisiä dokumentteja. (CSC – Tieteen tietotekniikan keskus Oy 2019)

Tämän opinnäytetyöprojektin kannalta kaikista keskeisin metatietotyyppi sekä tiedonhallinnan osa-alue on sisällöllinen metatieto. Sisällöllinen metatieto on nimensä mukaan tietoa, joka kuvaa tiedon sisältöä, sen tyyliä sekä kohderyhmää. Sisällön kuvaamisella tarkoitetaan esimerkiksi dokumenttiin liitettävien avainsanojen käyttöä, joilla voidaan luokitella dokumentti tiettyyn ryhmään tai kuvailla dokumentin tarkoitusta. Sisällöllinen metatieto voi kuvata myös kieltä, jolla tieto on eritetty sekä kyseiselle tiedolle määritettyjä parhaita käytötarkoituksia (Dalkir 2011, 276). (Dalkir 2011, 130.)

Sisällöllinen metatieto auttaa sen käyttäjiä päättämään, mistä tiedostossa on kyse sitä avaamatta. Se ei ole kuitenkaan sisällöllisen metatiedon päätarkoitus, joka on auttaa löytämään ja tunnistamaan tietoa (MerlinOne 2019; Smith 2007). Yksinkertaisimmillaan tiedon löytäminen sisällöllisen metatiedon avulla voi olla tiedostojen järjestämistä tiedostokansioissa samalla tavalla kuin hallinnollista metatietoa. Tämä vaatii kuitenkin sen, että sisällöllinen metatieto on määritelty tag-tiedoksi ja näkyy tiedostokansiossa tag-nimisessä sarakkeessa (kuva 5). Yleisesti ottaen sisällöllinen metatieto mahdollistaa tiedon löytämisen sanojen avulla paikallisesti sekä tietoverkkotasolla, eli internetissä (Cornell University Library 2002). (Kaario & Peltola 2008, 26.)



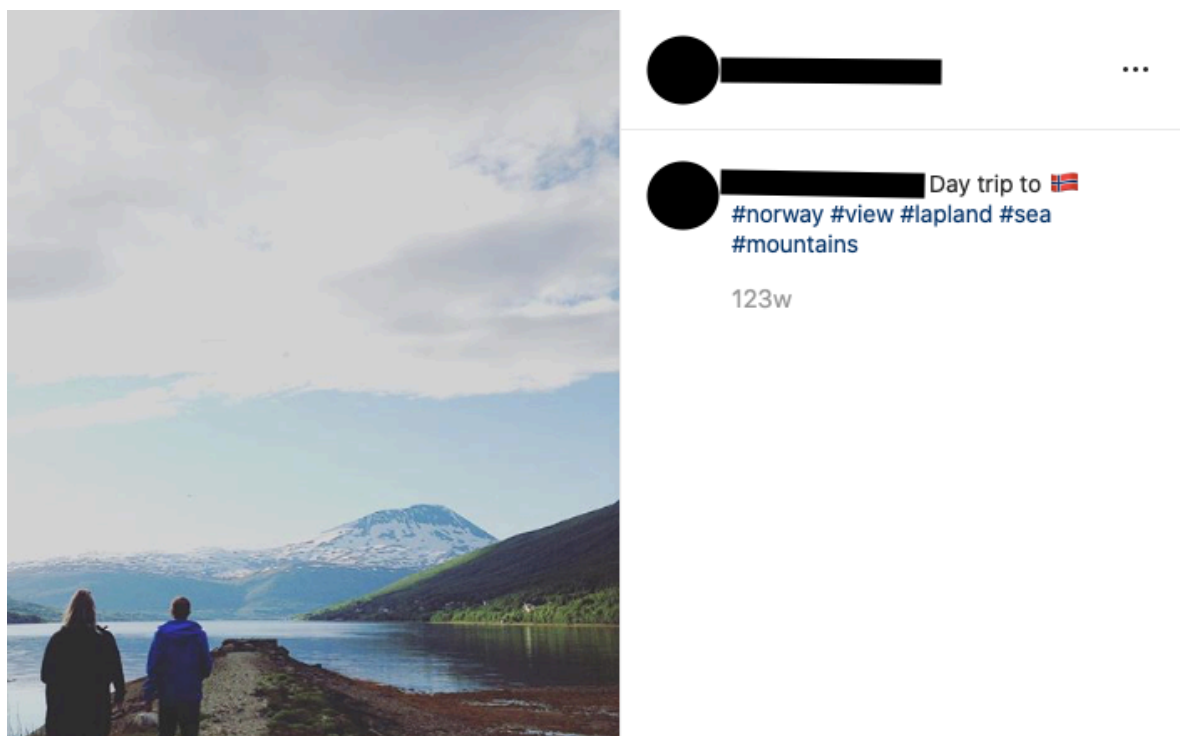
Kuva 5: Sisällöllinen metatieto tag-sarakkeessa

Vaikka sisällöllinen metatieto mahdollistaa tiedon etsimisen, sen käyttäminen ei ole aina viehättävää sen tuottamiseen liittyvistä syistä. Hallinnollisesta metatiedosta poiketen sisällöllisen metatiedon tuottaminen koetaan yleensä haastavaksi ja aikaa vieväksi, koska sen luominen on lähtökohtaisesti tiedoston luojaisten vastuulla. Sisällöllisen metatiedon automaattiseen generointiin dokumenteista on olemassa työkaluja, jotka helpottavat tätä luomisprosessia. Tällaisten ohjelmien avulla pystytään tunnistamaan relevantteja avainsanoja esimerkiksi dokumentin sisällöstä sekä nimestä, eikä dokumentin luoja tarvitse keksiä itse kaikkia dokumenttiin sopivia tunnisteita. (Kaario & Peltola 2008, 26-27; M-Files 2019a.)

### 2.1.2 Metatiedon neljäs laji

Tag tai suomenkielessä suositellumpi käyttösanana ”tunniste” tarkoittaa yksittäistä sanaa tai sanoista muodostuvaa yhdistelmää, joka on ensisijaisesti tietoa tiedostosta tai dokumentista (Sanastokeskus 2010). National Information Standards Organization (NISO), määrittelee tagit muun muassa tietona, joka kuvailee, selittää tai muuten tekee tiedosta helposti käytettävää ja hallittavaa. NISO:n ja Sanastokeskuksen määritelmät ovat hyvin lähellä metatiedon määritelmää, tietoa tiedosta. Tunnisteet ovatkin metatietoa, mutta ne erottuvat muusta metatiedosta niille määritellyillä ominaisuuksillaan. (Smith 2007.)

Muista metatiedoista eroten, tagit ovat vaikeammin kategorioitavia avainsanoja, joita tiedostojen luojat sekä käyttäjät voivat lisätä dokumentteihin. Niillä on huomattavasti laajempi käyttötarkoitus kuin millään muulla yksittäisellä metatiedon lajilla, sillä ne voivat kuvata tietoa kaikkien sen hallinnollisten, rakenteellisten ja sisällöllisten ominaisuuksien lisäksi myös tietoon liittyvien mielipiteiden, arviointien ja toimintojen perusteella. Tageja voidaan kutsua sosiaaliseksi metatiedoksi, sillä niillä on mahdollista tuoda tiedostojen käyttäjät osaksi tiedon kuvailua. Käyttäjät liittävät siis tunnisteita dokumentteihin ja ne näkyvät tiedoston metatiedoissa, kuten kuvassa 5. Monelle arkipäiväisempi esimerkki tunnisteista on Instagramin hashtagit (kuva 6). Nämä aihetunnisteet ovat myös tunnisteita, joita käyttäjät lisäävät julkaisuihinsa ja niitä voidaan käyttää julkaisujen etsimiseen (Sedu 2017). (Smith 2007.)



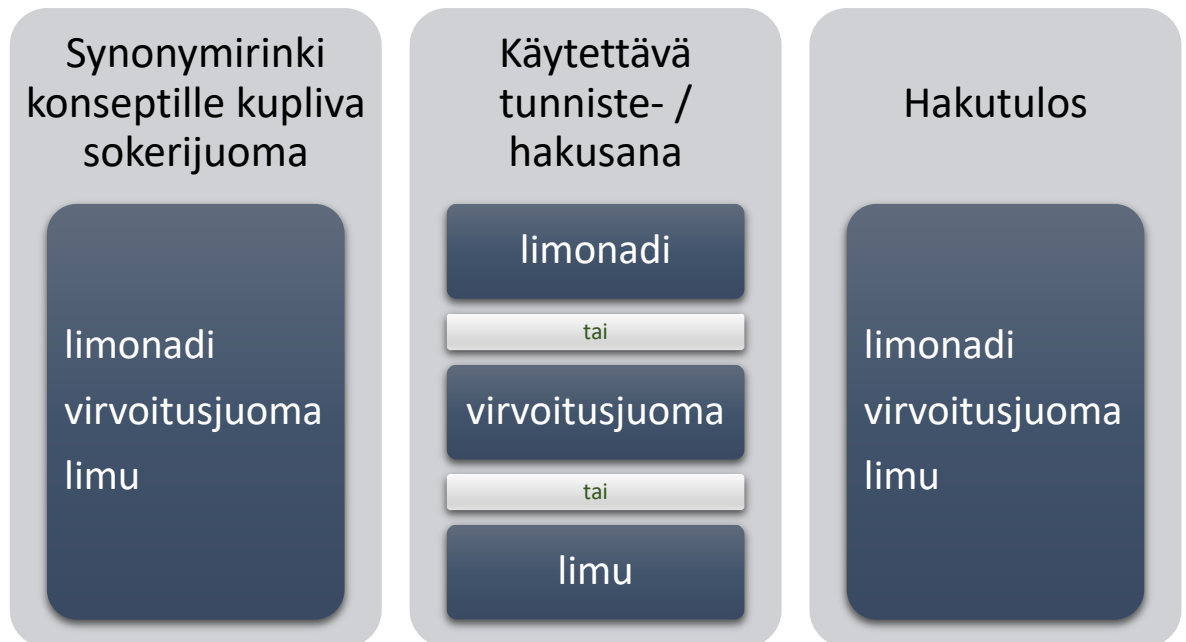
Kuva 6: Instagram-julkaisun hashtag-aihetunnisteet ovat tunnisteita (Kankaanpää 2017)

Tagien käyttötarkoituksista ja käytettävyydestä on käyty paljon väittelyjä, jotka ovat kohdistuneet erityisesti tunnisteiden yhteisöllisyyteen, erityispiirteisyyteen, hallintaan sekä asiantuntijuuteen. Sosiaalisuuden kohdalla on väitetty siitä, luovatko käyttäjät tunnisteita vain omaksi hyödykseen vai onko heidän tarkoituksena jakaa tietoa myös muille käyttäjille. Erityispiirteisyyden kohdalla keskustelua on ollut siitä, onko tunnisteiden hyvä olla täysin yksilöllisiä vai pitäisikö niitä standardoida tiedonhaun helpottamiseksi. Tagien hallinnan kohdalla väittelyä on taas käyty siitä, onko käyttäjille järkevää antaa vapaus tunnisteiden luomiseen vain olisiko esimerkiksi parempi antaa käyttäjille ehdotuksia tunnistesanojen käytön rajoittamiseksi. Asiantuntijuuden kohdalla keskustelu on liikkunut siinä, onko tagien painoarvo sama huolimatta käyttäjän osaamisesta. (Smith 2007.)

Cameron Marlow mukaan käyttäjien motivaatio tagien käyttämiseen määrittää millaisia tunnisteita he käyttävät (Smith 2007). Käyttäjä määrittelee siis itse käyttämiensä tunnisteiden yhteisöllisyyden, erityispiirteisyyden, hallisuuden sekä asiantuntijuuden. Organisaation sisällä tämä tarkoittaisi sitä, että organisaation tulee luoda yhteiset määritelmät, niin että kaikki käyttäjät toimisivat samalla tavalla. Tämä pätee kaikkien vapaasti määriteltävien metatietojen käyttöön, sillä niistä parhaimman hyödyn irti saaminen vaatii yhtenäistä kielen sekä sanastojen käyttöä ja niiden yhdenmukaista ymmärtämistä koko organisaation läpi (Kaario & Peltola 2008, 26-27).

Suurin haaste metatiedoissa käytettävän kielen yhdenmukaisessa käytössä ja ymmärtämisessä on se, että kieli on monimuotoista ja samasta asiasta voidaan käyttää useampaa eri sanaa jo yhdessä kielessä (Smith 2007). Haaste paisuu ymmärrettävästi silloin, kun organisaatiossa käytetään useampaa työkieltä, etenkin jos tunnisteissa käytettävää kieltä ja sanastoa ei ole rajattu tai sitä ei hallita millään tavalla.

Tunnisteiden hallintaan on luotu menetelmiä, joiden avulla pystytään kontrolloimaan käytettyjä tunnisteita, helpottamaan käyttäjien työtä. Yksi näistä tavoista on synonyymiringit (englanniksi synonym rings), joissa yhdistetään kaikki samaa asiaa merkitsevät sanat yhteen listaksi. Synonyymirinkiä käyttämällä tietokone osaa löytää kaikki dokumentit, joissa mainitaan mikä tahansa kyseisessä synonyymiringissä olevista sanoista (kuva 7). (Smith 2008, 67-69.)



Kuva 7: Synonymyymiringin toimintaperiaate

Toinen paljon käytetty tunnistesanojen kontrollointitapa on authority files, joka on muodoltaan kuin synonymyymiringit, mutta määrittelee ringistä yhden sanan suositelluksi käyttösanaaksi. Tämä on hyödyllistä esimerkiksi silloin, kun kyseessä on tuotteiden nimiä, jotka ovat saattaneet muuttua ajan myötä. Authority filen avulla tietokone ehdottaa käyttäjälle kaikista ajankohtaisinta sanaa ringistä, vaikka käyttäjä käyttäisi jotain toista synonymyymiringin sanaa. Esimerkiksi kuplivan sokerijuoman kohdalla tietokone antaa ensimmäisenä hakutulokseksi dokumentit, joissa on käytetty sanaa virvoitusjuoma, vaikka käyttäjä olisi etsinyt dokumenttia sanalla limonadi. (Smith 2008, 68-70.)

Nämä molemmat tavat helpottavat dokumenttien tunnistamista ja hakemista avainsanoilla silloin, kun organisaatiossa käytetään kahta tai useampaa sanaa samasta asiasta. Tällöin tunnisteiden luoja ei tarvitse olla varma, mitä sanaa muut ihmiset käyttävät samasta asiasta ja samalla tavalla dokumentin etsijän ei tarvitse tietää tarkasti, mitä samaa dokumentin luoja on käyttänyt dokumentin tunnistryhmässä. (Smith 2007.)

## 2.2 Tiedonhallinta

Metatiedot ja tunnisteet liittyvät tiiviisti tiedonhallintaan, mutta tiedonhallinta on myös muuta. Tiedonhallinta voidaan määritellä sanan käyttötarkoituksesta riippuen ainakin neljällä eri tavalla liiketoimintaympäristössä. Käsitteet tietokantojen hallinta, liiketoimintatietojen hallinta, tietämyksen hallinta sekä organisaation tiedon hallinta voidaan kaikki käsittää tällä yhdellä sanalla ja osaa näistä käsitteistä käytetään myös päällekkäin. Tästä listasta voidaan huomata, että nämä neljä mainittua käsitettä eivät ole rinnastettavissa toisiinsa täydellisesti, vaan ne jaottelevat tietoa. Tietokantojen hallinta kuvaa ainoastaan sellaista

tiedonhallintaa, jossa informaation sijaitsee, jonkin tietokannan sisällä. Samalla tavalla liiketoimintatietojen hallinnolla kattaa ainoastaan liiketoimintaan liittyvän informaation hallintaan, eikä pidä sisällään esimerkiksi organisaation muun datan hallintaa. Käsitteiden selkeiden eroavaisuuksien vuoksi näen, että tiedonhallinta kuin sateenvarjo, joka kokoaa alleen kaikenlaisen tiedon hallinnan, informaation muodosta ja hallintatavasta riippumatta. (Kaario & Peltola 2008, 3.)

Tämän opinnäytetyön kannalta tiedonhallinnan sateenvarjon alta kaikista merkittävistä käsitteistä on tietämyksen hallinta tai tietopääoman hallinta, joka laajentaa käsitteen tarkoitusta kaiken organisaation aineettomien omaisuususerien hallintaan. Tietopääoman hallinta tunnetaan suomenkielessä yleisesti englanninkielisellä termillä Knowledge management sekä lyhenteellä KM. Knowledge managementilla on pelkästään englanninkielessä yli kaksikymmentä hyväksyttyä määritelmää, jonka vuoksi niitä on jaettu käyttötarkoituksen mukaan erilaisiin konteksteihin. Näitä konteksteja ovat muun muassa liiketoimintaympäristö, kognitiotiede sekä prosessitekniikka. (Ståhle & Grönroos 1999, 11; Dalkir 2011, 5-6.)

Lähes kaikissa Kimiz Dalkirin (Dalkir 2011, 5-7) kokoamissa Knowledge managementin määritelmissä kiteytyy yksi asia. Knowledge management on toimintaa, jonka avulla yksilöllä oleva tieto tehdään käytettäväksi. Tiedon tekemisellä käytettäväksi tarkoittaa esimerkiksi hiljaisen tiedon siirtämisestä kommunikaation välityksellä ihmiseltä toiselle, yksilön hallitseman tiedon hyödyntämistä organisaation hyväksi tai teknologiaa ja prosesseja, joilla tietoa jaetaan organisaation jäsenille. (Kaario & Peltola 2008, 8; Dalkir 2011, 5.)

Tiedon tekeminen käytettäväksi teknologian avulla korostuu nykypäivänä tietokoneiden älykkyyden kasvaessa, etenkin tekstimuotoisen tiedon osalta. Tekstimuotoinen tieto ei kuitenkaan ole ainut tiedon muoto, jota voidaan hallita tietotekniikkaa hyödyntäen, sillä useita samoja toimintoja ja periaatteita voidaan käyttää myös äänen, numeroiden, kuvien ja minkä tahansa muun sähköisen informaation hallintaan. Näistä menetelmistä esimerkkejä ovat erilaiset dokumenttipankit, kuten kuvapankit ja asiakirjojen hallintajärjestelmät, sekä metatietojen hyödyntäminen, johon paneuduttiin edellisessä alaluvussa. (Kaario & Peltola 2008, 8 & 15.)

Tiedon käytettäväksi tekemisen ensimmäinen askel on tiedon taltioiminen. Se on myös osa tiedon elinkaarta, joka kuvaa millaisia erilaisia vaiheita tieto käy läpi sen ensiluomisen jälkeen. Tiedon elinkaaren vaiheet vaihtelevat tapauskohtaisesti paljonkin, muun muassa tiedon muodon mukaan. On kuitenkin pystytty tunnistamaan neljä sellaista päävaihetta, jotka kaikki tieto käy läpi elinkaarensa aikana (kuva 7). Tiedon taltioiminen on näistä vaiheista ensimmäinen ja tämän opinnäytetyön kannalta kaikista merkityksellisin. Taltiointi-



vaihteessa tiedolle määritellään suurimmaksi osaksi sitä kuvailevat tiedot, eli metatiedot, joita käsiteltiin aiemassa aliluvussa. (Kaario & Peltola 2008, 9-10.)

Tiedon taltioimisen jälkeen tiedon elinkaareissa tulee vastaan tiedon ylläpito ja hallinta, tiedon säilytys ja arkistointi sekä viimeisenä tiedon esittäminen, jakelu ja julkaisu. Tiedon ylläpitoa sekä hallintaa ja tiedon säilytystä sekä arkistointia ei varsinaisesti määritellä aikakuvassa peräkkäisiksi vaiheiksi, sillä ne voivat kumpikin esiintyä tiedon elinkaaren aikana useamman kerran ja niiden paikat saattavat vaihtua keskenään. Tiedon ylläpitäminen ja hallinta eivät koske tässä kohtaa ainoastaan yksittäisen dokumenttiin liittyviä toimenpiteitä, vaan myös organisaatiossa määriteltujen periaatteiden ylläpitämisen. Tämä tarkoittaa muun muassa sisältöön liittyvän luokittelun sekä arkistointi- ja tunnistetietojen ylläpitoa. Arkistointisääntöjen ylläpitäminen kuuluu sen sijaan tiedon säilytys ja arkistointivaiheeseen, tiedon säilyttämistä koskevien prosessien noudattamisen kanssa. (Kaario & Peltola 2008, 10-11.)



Kuva 8: Kaarion ja Peltolan kuvaus tiedon elinkaaren päävaiheista

Neljännessä luvussa esitelty tunnisteopas on luotu nimenomaan tiedon elinkaaren ensimmäisen kohdan tueksi, koska oppaan avulla dokumenttien luojat pystyvät helposti lisäämään taltiointivaiheessa tunnisteita dokumenttiin. Opas on myös osatiedon arkistointia, sillä opas sisältää tunnistetietojen rakennemäärittelyt ja se on siten yksi tunnistetietojen ylläpitotyökaluista.

### 3 Prosessin elinkaari ja taustamääritelmät

Tässä luvussa esitellään toimeksiannon prosessi sekä määritelmiä, joita tehtiin ennen oppaan laadintaa. Prosessin esittely sisältää määrittelyprosessin taustat ja tarpeen syntyminen sekä työvaiheiden kuvauksen alla kuvatun ajanjakson aikana (kuva 9).



Kuva 9: Prosessin ajanjakso ennen ja jälkeen toimeksiannon

Määrittelyprosessi sai alkunsa virallisesti toukokuussa 2019, jolloin aloitin tämän opinnäytetyö kirjoittamisen. Prosessin juuret löytyvät kuitenkin jo kesältä 2018, jolloin pohdiskelin kollegani pyynnöstä, millaisia avainsanoja olisi järkevää liittää erääseen dokumenttiin. Tuolloin kollegani ajatuksena oli luoda tunnisteet yhdelle kohdeyrityksen tärkeimmistä käyttöoppaista. Tämän pienen projektin aikana määrittelin aihetunnisteet jokaisen dokumentissa olevan otsikon alle, jotta dokumentista voisi hakea tarvittavia ohjeita tunnisteiden avulla. Projektin tavoite oli helpottaa otsikoihin merkityillä tunnistesanoilla ohjeiden löytämistä dokumentista ja tehdä lähes kaksisataasivuisesta oppaasta käyttäjäystävällisempi.

Yllä kuvattu projekti muokkaantui tämän opinnäytetyön toimeksiannoksi keväällä 2019, kun havaittiin, että yrityksen työntekijät eivät löytäneet tarvitsemaansa tietoa työpaikan tiedostopalvelimelta. Tämän havainnon myötä järjestin keväällä 2019 johdantoluvussa kuvattun kevyen haastattelukierroksen, joka vahvisti, että etenkin työaika- ja loma-asioita koskevat tiedot eivät olleet löydettävissä. Haastattelu paljasti myös, kuinka paljon tiedon etsimiseen on kulutettu aikaa ja kuinka useasti yrityksen työntekijät ovat käyneet keskeyttämässä kollegoidensa työtä selvittääkseen kaikille saatavissa olevaa tietoa (Haastattelut 02/2019).

Haastattelun pohjalta yritykselle päätettiin hankkia dokumenttienhallintajärjestelmä, jonka tueksi minun oli tarkoitus määritellä uudessa työkalussa käytettävien tunnisteiden rakennetta. Tutustuin touko-kesäkuussa 2019 erilaisiin dokumenttien hakutapoihin, ennen määrittelyjen tekemistä. Olen kuvannut erilaisia hakutapoja tämän luvun viimeisessä osiossa, josta löytyy myös tuki metatietoihin pohjautuvan dokumenttien hallinnan valitsemiselle.

Hakutapojen tutkimisen aikana aloitin Bluefors Oy:n tiedostopalvelimen ja tiedostojen tutkimisen. Tässä työvaiheessa oli tärkeää selvittää, millaista tietoa yrityksellä ylipäättensä on. Tiedon laadun ja muodon ymmärtäminen on tunnistesanojen luomisen perusta, koska metatieto, jota tunnisteet tai tagit ovat, on tietoa tiedosta (Smith 2007). Minun tarvitsi siis selvittää, millaista tietoa tunnisteilla tulee pystyä kuvaamaan, kun uutta dokumenttienhallintatyökalua käytetään. Nykypäiväisten tietosuojamääritysten vuoksi tunnisteiden laadun ja sisällön tutkiminen oli haastavaa, koska minulla ei ollut tarpeeksi oikeuksia kaikkien dokumenttien näkemiseksi.

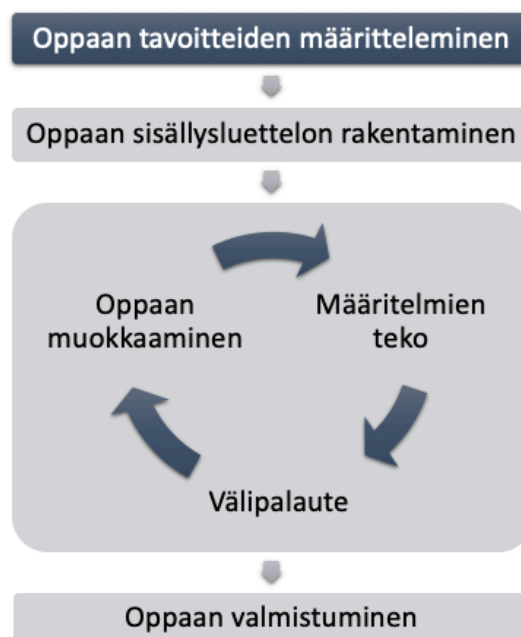
Toimeksiannon kannalta oli hyvä, että olen työskennellyt eri osastojen dokumentaation parissa lähes kaksi vuotta ja olen saanut sinä aikana hyvän yleiskäsityksen, millaisia dokumentteja yrityksessä luodaan. Vuonna 2018 olin päävastuussa yrityksen uusien dokumenttipohjien käyttöönotosta ja niiden kehittämisestä, joka lisäsi tietämystäni dokumenttien eri käyttötarkoituksista. Tuon projektin aikana minun tarvitsi luoda dokumenttipohjia erilaisiin julkaisutarkoituksiin, jolloin kartoitin dokumenttien ominaisuuksia julkaisemisen näkökulmasta. Samalla selvisi, mitkä osastot luovat dokumentteja kohdeyrityksessä ja julkaisukanavissa tiedostoja julkaistaan.

Tiedostopalvelimelta löytyvien tiedostojen sisällön ja laadun ymmärtäminen antoi minulle hyvää yleiskäsitystä siitä, millaisia käytössä olevat tunnisteet tulisivat olemaan tulevaisuudessa. Kirjasin tiedostojen tutkimisen aikana ylös myös dokumentteihin sopivia tunnistesanoja listaksi. Käytin tuota listaa hyödyksi tunnisteoppaan tunnistesananastossa sekä tun-

nistesanojen määrittelyssä. Sain listasta hyvän ymmärryksen, millaisia erilaisia ominaisuuksia tunnisteilla voi olla, joka auttoi minua kehittämään yhdenmukaisia määritelmiä tunnisteiden rakenteelle.

Tunnisteiden rakenteen ja niissä käytettävän kielen määrittelemisen tapahtui suurin piirtein saamaan aikaan kuin Bluefors Oy:n tiedostopalvelimelta löytyvien tiedostojen tutkiminen. Määritelmät keskittyivät kielen, merkistön, tunnisteissa käytettävien sanojen oletusmuodon sekä tunnisteiden yhdistelemistapojen valintaan. Nämä päätökset on esitelty tarkemmin seuraavassa luvussa, joka käsittelee tunnisteoppaan sisältöä. Lokakuussa 2019 valmistuneen tunnisteoppaan laatimisprosessin työvaiheet on kuvailtu erikseen seuraavassa alaluvussa sekä kuvassa 10.

### 3.1 Oppaan laatiminen



Kuva 10: Oppaan laatimisprosessin vaiheet

Guidelines for Tagging-oppaan ensimmäisen version kirjoittaminen alkoi heinäkuussa 2019. Tuohon aikaan olin jo ehtinyt ymmärtämään dokumenttien hakutapojen erilaiset ominaisuudet sekä selvittämään pääpiirteisesti, millaisia tiedostoja tuleva dokumenttienhallintatyökalu tulisi käsittelemään. Halusin, että opas sisältää ainoastaan käyttäjälle kaikista välttämättömimmän tiedon ja olisi niin helppolukuinen ja ytimekäs, että käyttäjät viitsisivät käyttää sitä. Määrittelin oppaan tavoitteiksi helppolukuisuuden sekä kompaktiuden. Näiden tavoitteiden pohjalta aloin hahmottelemaan oppaalle sisällysluetteloa.

Sisällysluettelo syntyi melko nopeasti ja se piti joidenkin lukujen järjestystä lukuun ottamatta saman rakenteen koko oppaan rakentumisen ajan. Oppaan rakentuminen tapahtui

jonkin aikaa päällekkäin tunnisteiden rakenteen määrittämisen kanssa, jonka vuoksi sisällysluettelo eli jonkin verran. Määritelmien muodostuminen oppaan kirjoittamisen kanssa päällekkäin aiheutti myös sen, että osa sisällöstä muuttui määritelmien syntyessä sekä muuttuessa. Esimerkiksi kieleen liittyvä osa oppaassa muuttui useamman kerran oppaan muodostumisen aikana, koska kielivalinta oli vaikea tehdä ja kieleen liittyviä ominaisuuksia tuli prosessin edetessä enemmän esille.

Käytin tunnistesanojen rakenteen määrittämiseen avukseni internetistä löytyviä artikkeleita sekä e-kirjoja. Sain myös kollegaltani, joka toimi tunnisteoppaan toimeksiantajan, tukea rakenteellisten määritysten tekemiseen. Palautin oppaan ensimmäisen version hänelle heinäkuussa, jonka jälkeen muokkasin opasta. Oppaan kirjoittaminen muodostui tässä vaiheessa kehämäiseksi prosessiksi, kuten kuvassa 10. Sain palautetta työstä, jonka myötä muutin määritelmiä ja oppaan sisältöä. Tämä toistui useamman kerran, mutta välillä välipalaute puuttui välistä, kun vain jokin määritelmä päivittyi.

Tein viimeiset päivitykset oppaaseen lokakuussa 2019, jolloin myös palautin lopullisen version kommentoitavaksi kollegalleni. Sain hyväksynnän valmiille oppaalle toimeksiantajalta syyskuun alussa 2019.

### **3.2 Hakutavan ymmärtäminen**

Ennen oppaan laatimista, määrittelyjen tekemistä sekä kohdeyrityksen tiedostopalvelimen tutkimista minun oli selvitettävä, millaisia dokumenttienhallintatapoja on olemassa. Vaikka projektille oli määritetty jo toimeksiannon yhteydessä siinä käytettävä dokumenttien hakutapa, mielestäni oli tärkeää tutkia erilaisia vaihtoehtoja sekä ymmärtää, miksi tunnisteisiin pohjautuva tapa oli valittu Bluefors Oy:n dokumenttienhallintaan.

Hakutavoilla tarkoitetaan dokumenttien hallinnassa erilaisia teknisiä ratkaisuja, jolla tiedostoja voidaan löytää niiden säilytykseen käytetystä sähköisestä varastosta, Bluefors Oy:n kohdalla tiedostopalvelimelta. Tämä tekninen ratkaisu on todella olennaista dokumentaationhallinnassa, sillä se määrittelee millaisilla ehdoilla ja miten laajasti dokumentteja voidaan hakea ja löytää. (IITC 2019.)

Dokumenttien eri hakutapoja ovat muun muassa tiedostojen kansiorakenne eli varastointilogiikka, dokumentin nimeämistapa, dokumentin sisältö sekä tiedoston metatiedoissa käytettävät tunnisteet. Jokainen näistä tavoista sopii erilaisiin tilanteisiin ja niillä on yksilöllisiä ominaisuuksia, vaikkakin jostain niistä löytyy myös yhtäläisyyksiä (IITC 2019).

Kohdeyrityksen dokumenttien hakutapa on perustunut alusta alkaen tiedostojen varastointilogiikkaan sekä väljästi tiedostojen nimeämiseen, joka selviää yrityksen tiedostopalvelinta ja dokumentteja tutkimalla. Nimeämisperiaatteet dokumenteissa sekä kansiorakenteissakin ovat perustuneet ainoastaan kunkin käyttäjän tai tiimin omiin näkemyksiin, eivätkä ole johdonmukaisia koko organisaation läpi. (Bluefors 2019e.)

### 3.2.1 Erilaiset hakutavat

Huolella tehty ja hyvin ylläpidetty kansiorakenne voi olla monessa toimintaympäristössä toimiva dokumenttien hakutapana, erityisesti sen visuaalisuuden vuoksi. Käyttäjän on helppo nähdä kansiorakenteesta yhdellä vilkaisulla, millaisiin dokumenttirakenteisiin tiedostot jakautuvat (kuva 11). Kansiorakenteen perusteella voidaan helposti päätellä, mitkä dokumentit ovat keskenään samanlaisia tai mitkä dokumentit liittyvät samaan aiheeseen. Yhden kansion sisällä voidaan esimerkiksi säilöä kaikkia organisaation videoita tai kaikkia yhteen markkinointiprojektiin kuuluvia dokumentteja lajiteltuna alakansioihin. (IITC 2019.)



Kuva 11: Käyttäjä näkee dokumenttirakenteet helposti kansiorakenteesta

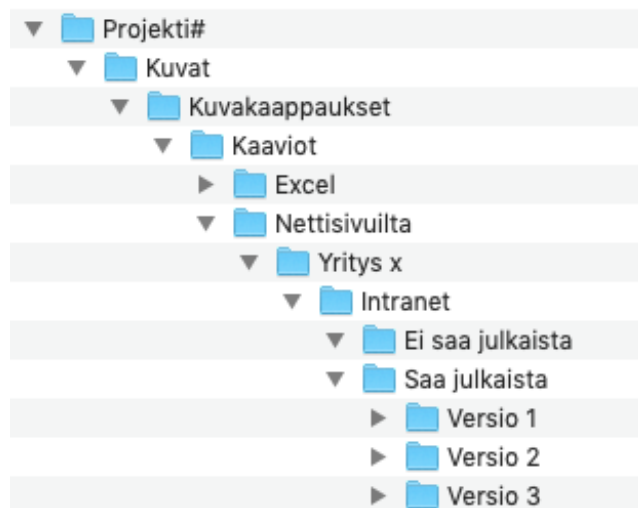
Kansiorakenteen huono puoli on kuitenkin siinä, että jonkun tarvitsee aina tietää, mitä kussakin kansiossa säilytetään. Tämä ei ole välttämättä haaste silloin, kun tiedostopalvelimella on harvoja käyttäjiä, mutta jos usealla kymmenellä organisaation jäsenellä on oikeus luoda ja muokata tiedostoja sekä niiden säilyttämiseen käytettäviä kansioita, muutoksissa on lähes mahdotonta pysyä perillä. Pelkästään kansioden toisistaan eroavat nimeämistavat tekevät kansiorakenteella tiedostojen hallitsemisesta hankalaa ja ajan myötä kaoottista (M-files 2019b).

Toinen tiedostojen varastointilogiikkaan kohdistuva haaste tulee esille tilanteissa, jossa samoja dokumentteja käytetään useampaan projektiin ja dokumentit halutaan säilyttää projektikohtaisesti. Tällainen tilanne pakottaisi joko useassa projektissa käytetyn dokumentin sijoittamista pois muiden projektitiedostojen yhteydestä tai dokumentin kopioimista useampaan kansioon (UC Merced Library 2019). Dokumenttien kopioiminen on hyvin

epätoivottua tiedostopalvelimen tilan käyttäminen sekä dokumentin päivittämisen näkökulmasta. Turhat dokumentit pakottavat organisaatioita käyttämään rahaa palvelimen muistitilan kasvattamiseen ja useassa paikassa sijaitsevien dokumenttien päivittäminen vaatii jokaisen kopion etsimistä päivittämistä varten.

Kolmas haaste dokumenttien kansiorakenteisiin perustuvassa haussa on Bluefors Oy:n työntekijän kanssa käydyssä keskustelussa esiin tullut kansiohierarkian logiikka. Syyskuun alussa käydyssä keskustelussa kävi ilmi, että erään osaston kansioissa hierarkia, eli tapa, jolla tiedostot on jaettu alikansioihin, esti tiedostojen hakemista ryhmänä. Dokumentit oli jaettu luomisvuoden mukaan kansioihin, eikä projektinumeroiden mukaan, joka olisi ollut toisen osaston työntekijöille tarpeellista. (Keskustelu 9.9.2019)

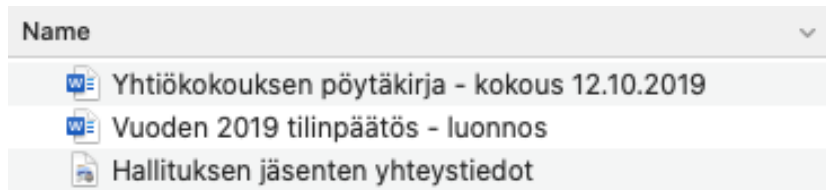
Hierarkian toinen haaste näkyy alla olevassa kuvassa (kuva 12), jossa kansiohierarkia on kasvanut niin pitkäksi, että alikansioiden alikansioista dokumenttien löytäminen on epäkäytännöllistä. Tällaisia kansiohierarkioita voi muodostua esimerkiksi silloin, kun tiedostojen määrä kasvaa ja kansioiden näkymää ensi silmäyksellä halutaan pitää siistinä.



Kuva 12: Esimerkki kansiohierarkian kasvamisesta liian pitkäksi

Dokumenttien nimeäminen on hyvin suosittu dokumenttien hallinnan hakutapa, jota käytetään monesti kansiorakenteen rinnalla. Se ei ole oikeastaan mikään ihme, sillä dokumentin nimeen voidaan helposti liittää tietoa dokumentista sekä sen sisällöstä (kuva 13). Dokumentteja on myös helppo etsiä niiden nimen perusteella, niin kauan kuin hakijalle on selvää, mitä nimeämiskäytäntöä kyseisessä dokumentissa on käytetty. Organisaatiossa, jossa tiedostoille pystytään suunnittelemaan sekä luomaan yhtenäinen dokumenttien nimeämismalli, tiedostojen yhtenäinen nimeäminen voi olla hyvä ratkaisu dokumenttien

hallintaan. Tähän tapaan pohjautuvat hakutapa edellyttää myös melko pientä tiedostomäärää, koska tiedon lisääntyessä tietokoneella kestää pidempi aika käydä läpi ja löytää kaikki hakukriteeriä vastaavat dokumentit.



Kuva 13: Tiedoston nimi kertoo tietoa dokumentista ja sen sisällöstä.

Tiedostojen nimeämis- ja varastointilogiikan lisäksi olemassa on kaksi hieman erilaista dokumenttien hakutapaa. Nämä ovat tiedoston sisältöön perustuva hakutapa sekä tunnisteisiin, eli dokumentin metatietoihin pohjautuvat hakutapa. Tiedoston sisältöön perustuva hakutapa on todella hyvä vaihtoehto silloin, kun käyttäjä tietää mitä haettava dokumentti pitää sisällään. Sisältöhaku kohdistuu ainoastaan dokumentin sisältöön ja poistaa käyttäjän kysymykset tiedoston sijainnista tai nimeämislogiikasta. Käyttäjän ei siis tarvitse tietää, missä hänen tarvitsema dokumentti sijaitsee, vaan ainoastaan mitä sanoja dokumentti sisältää. Tämä on kätevää esimerkiksi silloin, jos käyttäjä haluaa löytää kaikki dokumentit, joissa mainitaan tietty sana tai sanapari. (IITC 2019.)

Sisältöhaun huono puoli on kuitenkin se, että hakutulosten määrä voi olla valtava, eikä sisältöä voida luonnollisesti hakea dokumenteista, jotka eivät sisällä luettavaa kieltä (IITC 2019.). Tämä poissulkee pois hauista kaikki video-, kuva- ja äänitiedostot ja vähentää hakutavan arvoa. Pelkkä sisältöhaku ei siis ole hyvä ratkaisu ympäristössä, jossa käytetään paljon visuaalisia ja audiitiivisia tiedostoja, mutta yhdistämällä sen toisen hakutavan kanssa sisältöhaku voi olla oiva ratkaisu.

Viimeinen hakutapavaihtoehto perustuu metatietoihin, joita on käsitelty tarkemmin tämän raportin tietoperustassa. Metatietopohjaisessa tiedonhallinnassa dokumentin sijainnilla ei ole olennaista väliä tiedoston käyttäjälle, koska kuten dokumentin sisältöön pohjautuvassa toimintatavassa, myös tässä tavassa haku perustuu dokumentin sisältöön, sen sijainnin sijaan. Metatietoihin pohjautuvassa tiedonhaussa hakijan ei tarvitse kuitenkaan tietää, mitä sanoja dokumentissa on käytetty, vaan ainoastaan ymmärtää mihin dokumentti liittyy. Tämä johtuu siitä, että tunnisteita käytetään kuvaamaan dokumenttia ja sen sisältöä.

Tunnisteiden käyttäminen metatiedoissa mahdollistaa yrityksen tiedostojen varastoinnin vapaasti hakuprosessista erillään, jolloin turhien tiedostokopioiden muodostumiselta vältytään todennäköisemmin eikä dokumenttien käyttäjien tarvitse huolehtia kansiorakenteiden



ylläpidosta. Sisältöhakuun verrattuna metatietoihin pohjautuvan tiedonhallinnan vahvuus tulee esiin etenkin kuva-, video- ja äänitiedostojen kohdalla, koska myös näiden tiedostotyyppien metatietoihin on täysin mahdollista lisätä tunnisteita kuvailemaan tiedostoa.

### **3.2.2 Hakutavan valinta**

Hakutavan valinta ei ollut tässä projektissa minun vastuullani, mutta olen samaa mieltä päätöksestä toteuttaa kohdeyrityksen dokumenttien hallinta tunnisteiden avulla. Käytössä ollut kansiorakenteeseen ja pirstaleiseen tiedostojen nimeämislogiikkaan perustuva hakutapa on sekava, vaikea käyttää ja vaikea hallita. Tästä kertoo myös keväällä tehdyt haastattelut, joiden aikana tuli ilmi haasteet olemassa olevan tiedon löytämisessä. (haastattelut 02/2019).

Tiedostojen nimeämiseen perustuva hakutapa ei vaikuta teknisesti kovin toimivalta tavalla varastoida dokumentteja kohdeyrityksessä. Tämä johtuu siitä, että Bluefors Oy:llä on sen it-työntekijän kanssa käydyn keskustelun mukaan reilusti yli 300 000 dokumenttia sen tiedostopalvelimella ja oikean yksittäisen dokumentin löytäminen vie tietokoneelta pitkän ajan (keskustelu 16.9.2019). Kansiorakenteiden kohdalla tekninen vaikeus on käyttäjäoikeuksien määrittelyssä. Kaikilla työntekijöillä ei ole oikeuksia kaikkiin kansioihin, jonka vuoksi he eivät myöskään voi tallentaa luomiaan tiedostoja aina niihin kansioihin, joihin ne todellisuudessa kuuluisivat.

Dokumentin sisältöön perustuva hakutapa voisi olla toimiva tapa hallita Bluefors Oy:n tekstiä sisältäviä dokumentteja. Tämä ratkaisisi monen helmikuussa tehtyyn haastatteluun osallistuneiden haasteen löytää olemassa olevaa tietoa (haastattelut 02/2019). Hakutapa ei kuitenkaan toimisi kaikkien yrityksen dokumenttien kohdalla, sisältöhaun toimiessa vain dokumenttien kohdalla, joista voidaan hakea tekstiä. Kohdeyrityksessä käsitellään valtavaa määrää kuvatiedostoja päivittäin, muun muassa yrityksen tuotannossa. Jos nämä kaikki dokumentit jätettäisiin huomiotta, dokumenttienhallintatyökalusta ei olisi hyötyä monien kymmenien Bluefors Oy:n työntekijöiden työssä. Tunnisteita voidaan sen sijaan käyttää tiedoston muodosta riippumatta kaikissa sähköisissä dokumenteissa, jonka vuoksi se on parempi ratkaisu kohdeyrityksen dokumenttien hallintaan. Tunnisteiden käytön haaste on kuitenkin siinä, että jokaiselle dokumentille on määriteltävä tunnisteet ja organisaatiolla on oltava yhtenäinen tapa lisätä metatietoa dokumenteille. Tämä on yksi motiivi seuraavassa luvussa esitellyn Guidelines for Tagging -oppaan luomiselle.

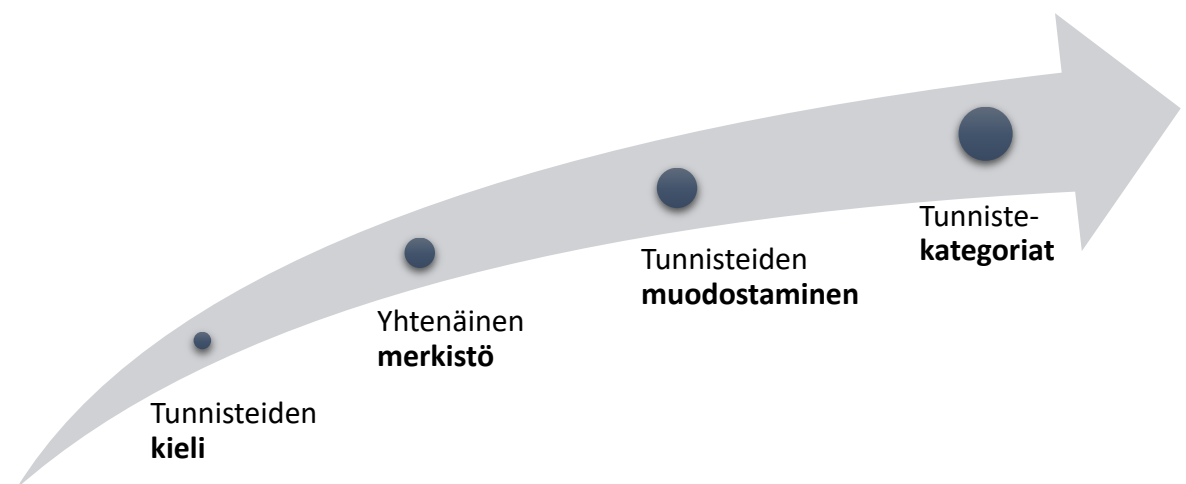
## 4 Tunnisteoppaan esittely

Tämä luku keskittyy kuvaamaan määrittelyprosessin aikana syntynyttä Guidelines for Tagging -tunnisteopasta, sen sisältöä sekä oppaan keskeisimpien määritelmien syntymistä. Lyhykäisyydessään tunnisteopas on Bluefors Oy:n työntekijöille tehty manuaali, jonka avulla he pystyvät luomaan itsenäisesti käyttökelpoisia tunnisteita luomilleen ja käyttämilleen dokumenteille. Tällä tavoin jokainen yrityksen työntekijä pystyy olemaan osa onnistunutta organisaation dokumenttienhallintaa.

Opas koostuu kolmesta pääluvusta sekä pikaoppaasta. Pääluvut jakavat tunnisteoppaan neljään osaan: tagien rakenne- ja kieliopas (Tag structure and definitions), tunnistesanojen käyttöopas (How to choose correct tags to apply to a document?), valmis tunnistesanalista (Existing tags) sekä pikaopas (Good tagging process in a nutshell). Tunnisteoppaan tarkoituksena on auttaa oppaan lukijaa muodostamaan kirjoitusasultaan yhdenmukaisia tunnisteita.

Tässä luvussa tarkastellaan erityisesti oppaan ensimmäistä päälukua sekä määritelmiä, jotka toimivat ohjeiden pohjana. Ensimmäisen pääluvun lisäksi käsitellään oppaan kolmannessa luvussa käsiteltäviä tunnistekategorioita. Muita oppaan lukuja ei käydä tässä luvussa läpi, sillä tämä opinnäytetyö keskittyy pääosin tunnisteiden kielelliseen rakentamiseen, jota käsitellään enimmäkseen oppaan ensimmäisessä pääluvussa Tag structure and definitions. Oppaan kaikki luvut ovat kuitenkin löydettävissä raportin lopussa olevasta oppaan sisällysluettelosta (liite 1).

Alla on kuvattu Guidelines for Tagging oppaan ensimmäisen pääluvun rakenne, johon on lisätty loppuun kolmannen luvun tunnistekategoriat (kuva 14). Tässä opinnäytetyöraportin luvussa aiheet mukailevat samaa aiherakennetta, joka lähtee suurimmasta konseptista kohti yksittäisen tunnisteiden luomista.



Kuva 14: Tunnisteoppaan aiherakenne

## 4.1 Tunnisteiden kieli

Gene Smithin on todennut, että suurin haaste metatiedoissa käytettävän kielen yhdenmukaisessa käytössä ja ymmärtämisessä on kielen monimuotoisuus ja sanojen semanttinen yhtäläisyys (Smith 2007). Sanojen merkitystä ei voida kuitenkaan määritellä ennen kuin tiedetään mitä kieltä käytetään. Mikäli tunnistesanojen käyttöön ei määriteltäisi yhtä yhtenäistä kieltä kaikkien käyttäjien käyttöön, tunnisteiden kielellinen monimuotoisuuden aiheuttama haaste on vaikea ratkaista. Tämän vuoksi Guidelines of Tagging -oppaan ensimmäisen määritelmä tunnisteiden kieli.

Toimintaympäristöissä on monesti määritelty virallinen käyttökieli, jota noudatetaan yhdenmukaisesti organisaation jokapäiväisestä sisäisestä kommunikaatiosta aina ulkoiseen viestintään sekä dokumentaationhallintaan. Bluefors Oy:ssä ei ole ennen tätä prosessia määritelty virallista työskentelykieltä, mutta henkilöstön kansainvälisyyden sekä yhteisen äidinkielen puuttumisen vuoksi, englanninkielestä on tullut luonnollinen käyttökieli kohdeyrityksessä. Bluefors Oy:ssä työskentelee kuitenkin sellaisia ihmisiä, jotka eivät puhu englantia. Tästä johtuen jotkin yleisistä dokumenteista, kuten Bluefors Oy:n Yhtiöpolitiikka sekä henkilöstötuihin liittyvät tiedostot, on toteutettu sekä englannin- että suomenkielellä. Tämä tarkoittaa sitä, ettei kaikki yrityksen viestintä ja dokumentointi tapahdu ainoastaan englanniksi. (Bluefors 2019d.)

Tunnisteiden hallittavuuden sekä yhtenäisyyden vuoksi tunnisteiden käyttöön oli kuitenkin määritettävä yksi kieli. Kielen valinta vei määrittelyprosessissa paljon aikaa, sillä aiheesta käytyjen keskustelujen pohjalta selvisi, että Blueforsin työntekijät olivat jokseenkin eri mieltä kielivalinnasta. Ainut järkevä vaihtoehto kohdeorganisaation dokumentaatiokieleksi oli kuitenkin englanti, koska se oli toiminut koko yrityksen 10-vuotisen toiminnan ajan epävirallisesti yrityksen toimintakielenä.

Pelkästään englanninkielen määrittelemine ei riitä, sillä englanninkielellä on useita variaatioita. Näissä jokaisessa variaatiossa osittain oma sanasto sekä erilaiset tavat kirjoittaa kielissä yhtenäisiä esiintyviä sanoja. Tästä esimerkkinä on brittienglanti sekä amerikanenglanti, joissa sadat jokapäiväisessä käytössä esiintyvät sanat ovat erilaisia kummassakin kielessä (Learning English 2017). Alla olevassa taulukossa on muutama esimerkki Bluefors Oy:ssä päivittäin käytetyistä sanoista, jotka eroavat amerikanenglannissa ja brittienglannissa toisistaan (Taulukko 1). Toisistaan täysin poikkeavien sanojen lisäksi brittienglannin ja amerikanenglannin kirjoitusasut poikkeavat useiden sanojen kohdalla toisistaan, jonka voi huomana alla olevasta taulukosta (Taulukko 2) (Oxford International Education Group 2019).

Taulukko 1: Sanaston eroavaisuuksia amerikanenglannin ja brittienglannin välillä (Beare 2019).

Amerikanenglanti	Brittienglanti	Suomennos
elevator	lift	hissi
mail	post	posti
schedule	timetable	aikataulu
truck	lorry	kuorma-auto
vacation	holiday	loma
wrench	spanner	jakoavain

Taulukko 2: Kirjoitusasun eroja amerikanenglannin ja brittienglannin välillä (British Council 2019).

Amerikanenglanti	Brittienglanti	Suomennos
alumin <u>u</u> m	alumin <u>i</u> um	alumiini
cance <u>l</u> ed	cancell <u>e</u> d	peruttu
center <u>u</u>	centr <u>e</u>	keskusta
color <u>u</u>	colour <u>u</u>	väri
fiber <u>u</u>	fibr <u>e</u>	kuitu
gray <u>u</u>	gre <u>y</u>	harmaa
license <u>u</u>	licenc <u>e</u>	lisenssi
met <u>e</u> r	metr <u>e</u>	metri
organiz <u>e</u>	organis <u>e</u>	järjestää

Kirjoitusasun erot ja eri sanaston käyttö ei muodostu välttämättä ongelmaksi silloin, kun puhutaan englanninkieltä tai lähetetään englanninkielisiä viestejä henkilöltä toiselle. Näistä eroavaisuuksista tulee kuitenkin haaste, jos halutaan, että jokainen käyttäjä käyttää yhtenäistä kieltä, esimerkiksi tunnistesanojen luomisessa. Tämän ongelman voisi ehkä ratkaista yhdistämällä eri englanninkielen variaatioiden samaa tarkoittavat sanat synonyymiringeiksi, joista puhuttiin tietoperustan. Itse näen tämän kuitenkin todella työlääksi vaihtoehdoksi, sillä jonkun tulisi käydä läpi kaikkien käyttöön haluttujen englanninkielen variaatioiden sanat läpi ja yhdistää eri tavoilla kirjoitetut vastineet toisiinsa. Tällainen projekti käsittäisi satojen tuhansia, ellei jopa yli miljoonan sanan läpikäymistä ja vaatisi ehdottomasti ammattilaisen osaamista.

#### 4.1.1 Kielen valinta

Tunnisteissa käytettävän kielen valinta oli hakutavan jälkeen merkittävin määrittelyprosessin kannalta tehty päätös. Jos kohdeyritys toimisi vain ja ainoastaan yhdellä kielellä ja

suurin osa sen työntekijöistä olisi äidinkieleltään yrityksen virallista kieltä osaavia, kielivalintaa tällaiselle organisaatiolle voisi olla melko suoraviivainen ja selkeä. Tässä tapauksessa valintaa ei voinut kuitenkaan tehdä yrityksen virallisen kielen perusteella, koska sitä ei oltu määritetty. Valintaa ei voitu perustaa myöskään työntekijöiden kansalaisuusprosenttiin, sillä se rajaisi ison osan suomen kieltä osaamattomista ihmisistä uuden dokumentinhallinnan ulkopuolelle. Samalla tavalla kuin hakutavan valinnassa, myös kielen valinnassa oli pidettävä huolta, että valinta mahdollistaa tunnisteiden käytön mahdollisimman laajasti ja mahdollisimman monelle työkalun käyttäjälle.

Käyttölaajuuden lisäksi kielivalinnassa oli otettava huomioon, miten yrityksessä käytetään kieltä ylipäättänsä. Mielestäni olisi erikoista määritellä millekään organisaatiolle useampia kieliä eri tarkoituksiin, ilman todella painavaa syytä. Kielivalintaa täytyy miettiä dokumentaationhallinnan lisäksi muiden yrityksen viestintäkanavien näkökulmasta. Näitä kanavia ovat esimerkiksi sisäiset työohjeet, markkinointi ja ulkoinen viestintä sekä toiminnanohjausjärjestelmä, jonka tietokannassa varastoidaan iso osa yrityksen tuotantoa koskevaa informaatiota. Minulle helpointa oli lähteä tutkimaan asiaa toiminnanohjausjärjestelmän kautta, sillä se on yksi tällä hetkellä eniten käyttämistä työkaluista ja sen käyttö ulottuu jokaiseen kohdeyrityksen osastoon.

Tutkiessani kohdeyrityksen Lean-toiminnanohjausjärjestelmästä löytyvää tietoa huomasin, että virallisen työkielen määrittelemättömyys näkyy vahvasti tuotantoon kohdistuvassa tiedossa. Leanissa esiintyy sekaisin niin amerikanenglantia kuin brittienglantiakin. Brittienglantia käytetään enemmän tuotteiden nimeämisessä, esimerkiksi alumiinisien osien nimissä, jotka oli useammin kirjoitettu brittityyliin ”aluminium”. Amerikanenglannin kirjoitusasu ”aluminum” oli kuitenkin käytetty myös alumiini-sanalle. Tämä oli mielestäni hieman huolestuttavaa, sillä toiminnanohjausjärjestelmän käyttäjältä jää tietoa löytämättä, mikäli hän ei osaa etsiä sitä aina kumpaakin kirjoitusasua käyttäen. (Bluefors 2019f)

Jotkin sanat Leanissa noudattavat yhdenmukaista kirjoitusasua läpi järjestelmän. Tällaisesta on esimerkkinä sana metri, joka kirjoitetaan amerikkalaisittain ”meter” kaikissa pituusmittayksiköissä. Myös toiminnanohjausjärjestelmästä löytyvissä ostotilauksissa käytetään yhdenmukaista kirjoitusasua peruutetuille tilauksille, joiden statuksessa lukee ”cancelled”. Tämän ei kuitenkaan noudata amerikanenglannin kirjoitusasua, joka olisi ”cancelled”, yhdellä L-kirjaimella. Meter- ja cancelled-sanojen yhdenmukaisuuden salaisuus löytyy kuitenkin siitä, että kumpikin tieto on määritetty ohjelmaan, eikä käyttäjät pysty muokkaamaan sitä, toisin kuin tuotteiden nimet, jotka ovat käyttäjien luomaa vapaata tekstiä. (Bluefors 2019f)

Toiminnanohjausjärjestelmän jälkeen tutkin kohdeyrityksen työntekijöiden tuottamia dokumentteja. Dokumentin kirjoittajasta riippuen sanoja oli kirjoitettu joko amerikanenglantia tai brittienglantia hyödyntäen. Joissakin dokumenteissa oli myös käytetty kumpaa variaatiota sekaisin. Dokumenttien tarkastelussa tuli vastaan myös uusi haaste, yhdyssanat, joita käsitellään tarkemmin tämän luvun viimeisessä aliluvussa. (Bluefors 2019e)

En pystynyt tekemään päätöstä kielivalinnasta olemassa olevien dokumenttien enkä toiminnanohjausjärjestelmässä olevan tiedon avulla, jonka vuoksi kävin keskustelemassa aiheesta kohdeyrityksen markkinointiasiantuntijan sekä toimistoassistentin kanssa. Kummassakin keskustelussa painotettiin, että yrityksen suunta on kohti Yhdysvaltoja, muun muassa sinne avatun tytäryhtiön vuoksi (Bluefors 2019b). Tulevaisuuden suunta on mielestäni hyvä syy kannattaa amerikanenglantia viralliseksi kieleksi. Tätä valintaa puolsi myös toimistoassistentin esille tuoma huomio kirjoitusasusta. Toimistoassistentin mukaan amerikanenglannin kirjoitusasu on helpompi kuin brittienglannissa. Tämä näkyy sanoissa kuten väri (color) ja keskusta (center), jotka kirjoitetaan amerikanenglannissa samalla tavalla kuin ne lausutaan. Brittienglannissa sanoihin lisätään kirjaimia tai niiden paikat muuttuvat (taulukko 2). Helpompi kirjoitusasu voisi vähentää virheiden muodostumista ja siten ylimääräisten tunnisteiden syntymistä. (Keskustelu 12.6.2019.)

Olemassa olevien dokumenttien ja toiminnanohjausjärjestelmän tutkimisen sekä kohdeyrityksen työntekijöiden kanssa käytyjen keskustelujen perusteella voisi olettaa, että amerikanenglanti tulisi olemaan yrityksessä käytetympi kieli pitkällä tähtäimellä. Tämän oletuksen ja yrityksen tulevaisuuden laajenemissuunnan perusteella dokumentaationhallinnassa käytettäväksi kieleksi määriteltiin amerikanenglanti.

#### **4.1.2 Ei-englanninkielisissä dokumenteissa käytettävä tunnistekieli**

Kielivalinnan tekeminen sai minut huojentumaan sekä huolestumaan. Kuten aiemmin tässä raportissa mainittiin, osa Bluefors Oy:n sisäisistä dokumenteista kirjoitetaan suomeksi. Lisäksi kohdeyritykseltä löytyy dokumentteja muilla kielillä, koska osa asiakkaista vaatii myyntisopimuksia eri kielellä. Miten käyttäjät pystyisivät tunnistamaan nämä dokumentit tunnisteiden avulla, jos kaikki tunnisteet tulee olla kirjoitettuna amerikanenglannin kieliasua noudattaen?

Tähän loin ratkaisuksi kielitunnisteet, joista löytyy erimerkkitaulukko alta (Taulukko 3). Taulukossa on käytetty #-merkkiä kuvaamaan, että hashtagin jälkeen oleva sana on yksittäinen tunniste. Kuten taulukosta voi huomata kielitunnisteet on luotu tunnistamaan ei-englanninkieliset sekä monikieliset dokumentit niissä käytetyn kielen tai kielten perus-

teella. Kielitunniste koskee kaikkia muita tiedostoja paitsi täysin englanninkielisiä dokumentteja. Englanninkielisiin dokumentteihin ei tarvita tätä lisätunnistetta, sillä kohdeyrityksen oletustyöskentelykieli on englanti. On kuitenkin järkevää käyttää englanninkielen tunnustetta silloin, kun dokumentissa käytetään useampaa kieltä. Tällöin dokumentin käyttäjä tietää, että monikielisessä dokumentissa on mukana varmasti myös englanninkielistä tekstiä.

Taulukko 3: Esimerkkejä kielitunnisteista

<b>Tiedoston tekstikieli / -kielet suomeksi</b>	<b>Tiedoston tekstikieli / -kielet englanniksi</b>	<b>Tunniste(et)</b>
suomi	Finnish	#finnish
ranska	French	#french
hollanti	Dutch	#dutch
espanja ja englanti	Spanish and English	#spanish #english
suomi, ruotsi ja englanti	Finnish, Swedish and english	#finnish #swedish #english
englanti	English	

## 4.2 Merkistöstä yhtenäisyyttä tunnisteisiin

Guidelines for Tagging -oppaaseen on ollut kielen määrittelemisen lisäksi oleellista tämentää, millaista merkistöä tunnistesanojen luomisessa voi käyttää, joka tehtiin oppaaseen heti kielen määrittävän luvun jälkeen. Merkistöllä tarkoitetaan tässä yhteydessä aakkosnumeerisia merkkejä, eli kirjaimia ja numeroita, sekä kaikkia erikoismerkkejä ja symboleita, joita voi luoda tietokoneen näppäimistöstä löytyviä näppäimiä sekä näppäinyhdistelmiä käyttämällä. Merkistö on määritelty oppaaseen ensisijaisesti edesauttamaan yhdenmukaista merkkien käyttöä. Joissain tietokoneohjelmissa on myös joitain rajoituksia merkkien käytössä ja merkistöstä on pyritty tekemään oppaaseen mahdollisen monia ohjelmia palveleva, sillä dokumenttienhallinta työkalua ei ole valittu tässä vaiheessa.

Yhdenmukaisen merkkien käytön tarkoituksena on yhtenäistää tunnisteiden kirjoitusasua silloin kun kyseessä on yksinkertainen yhdestä sanasta muodostuva tunniste tai kun tunniste muodostuu sanojen, numeroiden tai sanojen ja numeroiden yhdistelmistä. Merkkien yhdenmukainen käyttö on tärkeää aivan jokaisen tagin kohdalla, sillä sen avulla ehkäistään useiden tunnistevariaatioiden syntymistä ja siten jälleen ylimääräisten tagien luomista. Tällaisia variaatioita on mahdollista syntyä etenkin silloin, kun sanan kirjoitusasussa on eri käytäntöjä esimerkiksi väliviivan käytössä, tai sanan yleistyneeseen kirjoitusasuun

kuuluu jokin erikoismerkki. Muita tapauksia ovat lyhenteissä tai erisnimissä käytetyt sanoja korvaavat erikoismerkit, kuten &-merkki, sekä erisnimissä käytetyt ei-englanninkielisistä aakkosista löytyvät aakkoset, kuten skandinaaviset kirjaimet Å, Ä ja Ö.

#### **4.2.1 Merkistön tutkiminen**

Koska tunnisteiden kieleksi on valittu englanti, tulisi tunnisteiden olla aina englanninkielisiä sanoja tai englanninkielisistä sanoista ja numeroista muodostuvia yhdistelmiä. Tämä ei kuitenkaan ole realistista, sillä osa kohdeyrityksen asiakkaista ja muista sidosryhmistä on organisaatioita, joiden nimet ovat jotain muuta kieltä kuin englantia. Lisäksi sidosryhmien edustajien sekä Bluefors Oy:n työntekijöiden nimistä voi olla ei-englanninkielisiä. Tästä johtuen monissa tunnisteissa, jotka sisältävät erisnimen tai erisnimiä, ei voida välttää ääkkösten, eli suomalaisista aakkosista löytyvien Å, Ä ja Ö-kirjaimien, tai muiden ei-englanninkielisistä aakkosista löytyvien kirjaimien esiintymiseltä. Näiden kirjainten salliminen tunnisteissa mahdollistaisi kuitenkin jälleen eri kirjoitusasuvariaatioiden muodostumista tunnisteissa, jolta halusin välttää tässä prosessissa.

Ei-englanninkielisistä aakkosista löytyvien kirjainten aiheuttamista kirjoitusasuvariaatioista löytyy esimerkkejä kohdeyrityksen Lean-toiminnanohjausjärjestelmän tietokannasta, jossa tällaisia kirjaimia sisältäviä erisnimiä on kirjoitettu kahdella eri tavalla. Konkreettinen esimerkki on Ü-kirjaimen sisältävä erisnimi, joka löytyi tietojärjestelmästä sekä Ü-kirjaimella että U-kirjaimella kirjoitettuna (Bluefors 2019f). Tämänkaltaisten tilanteiden vuoksi halusin määritellä tunnisteisiin käytettävät kirjaimet ja miten kirjaimia tarvittaessa korvataan, kun ne eivät ole osa määriteltä merkistöä. Määritetyt käyttökirjaimet sekä kirjainen korvaustapa on nähtävissä taulukoista 4, 5 ja 8.

Erikoismerkkien käytössä pätee paljon samat haasteet kuin ei-englanninkielisten aakkosten käytössä. Mikäli erikoiskirjaimien käytölle jätettäisiin täysin vapaat mahdollisuudet, tunnisteiden kirjoitusasua ja kirjoitusasuvariaatioita olisi todella vaikea hallita. Tämän vuoksi erikoismerkeistä on haluttu päästä mahdollisimman paljon eroon. Erikoismerkit eivät esiinny väliviivaa ja et-merkkiä lukuun ottamatta juuri koskaan erisnimissä tai muissa tunnisteisiin kohdistuvissa sanoissa, jonka vuoksi ne voisi pitkälti kieltää kokonaan käytöstä. Et-merkki, eli &-merkki, sekä väliviiva esiintyvät kuitenkin melko usein englanninkielisessä kirjoituksessa, jonka vuoksi niiden poisjättämistä ja mukana pitämistä täytyi miettiä hieman enemmän.

Kohdeyrityksen dokumentteja tutkiessani huomasin, että mahdollisuus väliviivan käyttöön on aiheuttanut erilaisia kirjoitusasuvariaatioita englanninkielisissä sanoissa. Esimerkiksi eräässä dokumentissa oli kirjoitettu sana kolmas osapuoli englanniksi väliviivalla ”third-



party” sekä ilman väliviivaa ”third party”. Third party kirjoitetaan Oxford dictionaryn mukaan ilman väliviivaa ja harkitsin tästä syystä väliviivan jättämistä pois kokonaan tunnisteisiin käytettävästä merkistöstä (Oxford Dictionaries 2019). Väliviivaa käytetään kuitenkin muissa Bluefors Oy:ssä päivittäin käyttämissä sanoissa, kuten sanassa pre-cooling, eli esijäähdytys. Väliviiva otettiin osaksi merkistöä tunnisteoppaan ensimmäisessä versiossa, koska se esiintyi kohdeyrityksen jokapäiväisessä sanastossa. Viimeisestä Guidelines for Tagging -oppaan versiosta väliviiva kuitenkin poistettiin merkistöstä, koska sen käytöstä olisi luultavammin aiheutunut erilaisia kirjoitusvariaatioita, kuten tämän kappaleen alussa mainitusta third party -sanasta.

Aioin korvata &-merkki oppaassa n-kirjaimella, mutta opasta koskevan Tarmo Hiltuselta saamani välipalautteen jälkeen et-merkki päätettiin korvata and-sanalla. Tämä johtui siitä, että N-kirjaimella korvattaessa yrityksessä päivittäin käytettävästä lyhenteestä R&D tulisi RND, joka tarkoittaa, monelle vastinetta sanalle random, eli satunnainen (Hiltunen 6.8.2019). &-merkin muuttaminen n-kirjaimeksi tai sen pois jättäminen kokonaan saattaisi muuttaa myös muiden sanojen tai lyhenteiden merkitystä. Tämän vuoksi et-merkin korvaaminen päätettiin määrittää and-sanaa käyttäen.

#### 4.2.2 Merkistön määritteleminen

Koska merkistöstä haluttiin tehdä tunnisteiden rakennetta mahdollisimman paljon määriteltävä, päätin supistaa sen niin minimiin. Tästä syystä kaikki erikoismerkit jätettiin pois, yhtä poikkeusta lukuun ottamatta. Tämä poikkeus on alaviiva, joka tarvitsi jättää merkistöön tunnisteiden muodostamista varten. Tunnisteiden muodostamiseen sekä niihin liittyviin pohdintoihin palataan tämän aliluvun jälkeen.

Alaviivaa lukuun ottamassa erikoismerkkien ja kaikkien muiden symbolien lisäksi merkistöstä jätettiin pois kaikki muut kuin englanninkieliset aakkoset sekä numerot arabialaiset numerot (taulukko 6). Tämä johti käytännössä siihen, että merkintöihin jätettiin vain alla olevissa taulukoissa luetellut merkit, eli aakkoset A-kirjaimesta Z-kirjaimeen, numeeriset merkit 0-9 sekä alaviiva (taulukko 4; taulukko 5; taulukko 6; taulukko 7). Kirjainmerkistöön on määritelty erikseen pien- ja suuraakkoset työnohjaajalta saamani välipalautteen jälkeen, jotta käyttäjät tietävät, että kumpaakin vaihtoehtoa voidaan käyttää tunnisteissa, eivätkä ne aiheuta erilaisia kirjoitusasuja, mikäli kirjainten arvo määritellään tulevassa työkalussa samaksi (Hiltunen 6.8.2019).

Taulukko 4: Tunnisteisiin hyväksytyt suuraakkoset

A	B	C	D	E	F	G	H	I	J
---	---	---	---	---	---	---	---	---	---

K	L	M	N	O	P	Q	R	S	T
U	V	W	X	Y	Z				

Taulukko 5: Tunnisteisiin hyväksytyt pienaakkoset

a	b	c	d	e	f	g	h	i	j
k	l	m	n	o	p	q	r	s	t
u	v	w	x	y	z				

Taulukko 6: Tunnisteisiin hyväksytyt numerot

0	1	2	3	4	5	6	7	8	9
---	---	---	---	---	---	---	---	---	---

Taulukko 7: Ainut tunnisteisiin hyväksytty erikoismerkki

—
---

Edellä tunnistesanoihin hyväksytyt kirjaimet sisältävät ainoastaan englanninkielisiä kirjaimia, jonka vuoksi niihin kuulumattomien kirjainten korvaamiseen oli tehtävä ohjeistus. Tämän ohjeistuksen tekeminen oli tärkeää, sillä erikoisaakkosten korvaamisesta on olemassa useampi erilainen ohjeistus lähteestä riippuen. Esimerkiksi Suomen passissa alareunassa olevassa koneluettavassa tekstissä Å-, Ä- ja Ö-kirjaimet on korvattu AA-, AE- ja OE-kirjaimilla. Myös Nina Alajoki Kokkola ehdottaa käyttämään tätä samaa kirjoitustyyliä (Vähäsarja 2015). Tätä vastoin yksi suomen suurimmista matkatoimistoista kehottaa matkustajia korvaamaan sanaiset kirjaimet Intian viisumihakemuksiin A-, A- ja O-kirjaimilla (Aurinkomatkat). Erilaisten virallisten näkemysten vuoksi ja erityisesti useiden eri kirjoitusasuvariaatioiden välttämiseksi Guidelines for Tagging -oppaaseen on määritelty yksi tapa korvaamaan ei-englanninkielisiä aakkosia (taulukko 8).

Valinta tunnisteissa käytettävään erikoisaakkosten korvaamistapaan valittiin täysin puhtaasti luettavuuden sekä yksinkertaisuuden näkökulmasta. Näiden kahden periaatteen pohjalta päätin, että erikoisaakkoset korvataan aina mahdollisuuksien mukaan sitä eniten muistuttavalla yksittäisellä kirjaimella tai tarvittaessa useammalla kirjaimella, joista löytyy esimerkit alta (taulukko 8).

Taulukko 8: Esimerkkejä erikoisaakkosten korvaamiseen tunnisteissa

Kirjain	Å	Ä	Ö	Ü	Ø	Æ	Ç
Korvaaja	A	A	O	U	O	AE	C

### 4.3 Tunnisteiden muodostaminen

Tämän alaluvun sisällä esitellään prosessin aikana määritelty tapa muodostaa tunniste-sanoja yhtenäistä kirjoitusasua noudattaen. Kirjoitusasulla tarkoitetaan tässä kontekstissa sitä, millä tavoin edellisessä kappaleessa määriteltyjä merkkejä yhdistellään toisiinsa yhdenmukaisiksi tunnisteiksi. Oppaaseen mietittiin määrittelyiksi esimerkiksi pien- ja suuraakkosten sekä sanojen yksikkö- ja monikkomuotojen käyttöä. Tärkein kaikista tunnisteiden muodostamiseen liittyvistä määrittelyistä on kuitenkin tapa, jolla useampia sanoja yhdistetään yksittäiseksi tunnisteeksi. Kaikki nämä määrittelyt on tehty jälleen tukemaan yhdenmukaisuutta tunnistesanojen rakenteessa sekä rajoittamaan erilaisten kirjoitusasuvariaatioiden muodostumista. Alla olevassa taulukossa on kuvattu, miten erilaisia kirjoitusasuvariaatioita yhdestä tunnisteesta voisi syntyä, mikäli tageille ei määriteltäisi yhtenäistä kieltä, merkistöä ja muodostamistapaa (taulukko 9). Erot ensimmäiseen vaihtoehtoon on merkitty oranssilla värillä.

Taulukko 9: kahdeksan esimerkkiä yhden tunnisteiden kirjoitusasuvariaatioista

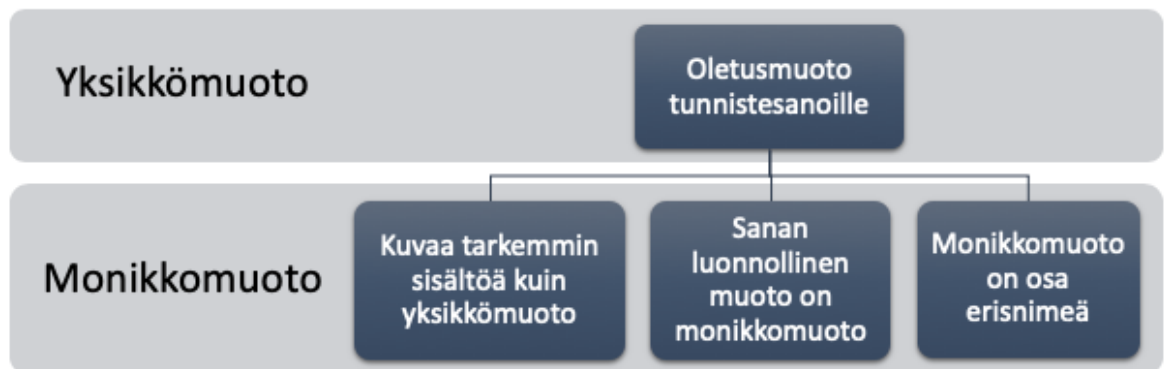
aluminumservicecase	aluminum_service_case
aluminumservicecases	aluminum_service_cases
Aluminiumservicecase	aluminium_service_case
aluminiumservicecases	aluminium_service_cases

Oppaan ensimmäisessä versiossa pien- ja suuraakkosten käyttö oli rajattu niin, että käyttäjät voisivat käyttää ainoastaan pienaakkosia. Tässä vaiheessa minulla oli käsitys, että pien- ja suuraakkoset ovat eriarvoisia ja vaikuttavat kirjoitusasuvariaatioiden syntymiseen. Välipalautteen perusteella suur- ja pienaakkosia koskeva määräitys jätettiin kuitenkin pois viimeisimmästä versiosta. Ainoat Guidelines for Tagging -oppaaseen jätetyt tunnisteiden kirjoitusasua ja muodostamista määrittelevät kohdat käsittelevät yksiköiden ja moniköiden käyttöä sekä sanojen yhdistämistä tunnisteeksi. Nämä aiheet on ohjeistettu Guidelines for Tagging-oppaan alaluvussa ”Spelling”.

Näin yksikkö- ja monikkomuotojen käytön määrittelyn merkittäväksi oppaaseen, koska sen avulla voidaan ehkäistä virheellisten tunnisteiden käyttöä sekä ylimääräisten tunnisteiden luomista. Kuten taulukosta 9 voi nähdä, jopa puolet kirjoitusvariaatioista on luotu ainoastaan vaihtamalla tunnisteiden monikkumuotoa.

Tunnisteiden oletusmuodoksi on määritelty tunnisteoppaaseen yksikkömuoto. Yksikkömuoto oletuksena on perusteltu Guidelines for Tagging -oppaassa sillä, että monikkumuodon käyttäminen on monesti harhaanjohtavaa, eikä kuvaa kyseistä dokumenttia oikein.

Silloin kun dokumentti sisältää vain yhden kuvan, tunniste sana "kuvat" kuvailisi tiedostoa väärin. Jos tiedosto on kollaasi useammista kuvista, kuvat-tunniste on kuitenkin oikea vaihtoehto. Tämä on yksi kolmesta poikkeuksesta, jolloin yksikkömuodon voi vaihtaa monikkomuotoon. Tagien yksikkö- ja monikkumuotojen käyttö on kuvattu alla olevassa kuvassa (kuva 15).



Kuva 15: Yksikkö- ja monikkumuotojen käyttö tunnisteissa

Ensimmäinen perusteltu perustelu monikkumuodon käyttöön on tilanne, jossa monikko kuvaa dokumentin sisältöä oikeammin kuin sanan yksikkö. Tiedosto voi esimerkiksi sisältää useita ohjeita yhden ohjeen sijasta, jolloin tunniste "ohjeet" kuvaa tiedostoa oikeammin kuin tunniste "ohje".

Toinen poikkeus perusteltu syy monikkumuodon käytölle on sanan luonnollinen monikko-muoto. Tästä ovat hyviä esimerkkejä sanat "sakset" (englanniksi scissors) ja "housut" (englanniksi pants). Saksien käytöstä opastavaan turvaohjeistuksen metatietoihin tulee tunniste "sakset" ja työhousujen kokotaulukoon liitetään tunniste "työhousut", koska sanat ovat luonnollisesti monikkumuotoisia.

Viimeinen perustelu monikon käyttöön on monikko erisnimessä. Joissakin yritysten nimissä esiintyy virallisesti sana monikkumuodossa, eikä tätä voida muuttaa tunnisteissa. Yrityksen nimen muuttaminen tunnisteessa voisi hypoteettisesti jopa aiheuttaa kahden eri yrityksen sekoittumisen keskenään. Esimerkiksi Paristo Oy ja Paristot Oy saisivat yhteisen tunnisteeseen paristo\_oy ja siten dokumentit voisivat mennä sekaisin näiden yritysten välillä.

Tunnisteoppaan Spelling-luvun toinen osio käsittelee tunnisteissa käytettävien sanojen yhdistelemistä yhdeksi tunnisteeksi. Piling words and characters into a single tag -alaluvun avulla käyttäjät voivat muodostaa yhdestä sanasta rakentuvien tunnisteiden sijaa useammista sanoista, numeroista tai sanoista ja numeroista muodostuvia tunnisteita. Tunnis-

teiden yhdistämiseen ei olisi välttämättä huomattavaa tarvetta, jos tunnisteissa käytettävässä kielessä on tyypillistä kirjoittaa yhdyssanat yhteen. Esimerkiksi suomenkielessä on mahdollista kirjoittaa yhtä konseptia kuvaava sana putkeen ilman välilyöntejä (taulukko 10). Prosessin aikana kohdeyrityksen tunnistekieleksi valittiin englanti, jossa suurin osa yhdyssanoista kirjoitetaan joko sanat peräkkäin sanaväleillä erotellen tai of-sanalla ja sanaväleillä sanat yhteen liittäen. Alla olevassa taulukossa on kuvattu suomen- ja englanninkielen välinen ero, yhden konseptin kuvaamisesta mahdollisimman lyhyesti (taulukko 10).

Taulukko 10: Vertailu yksittäisten konseptien kirjoitustavoista suomeksi ja englanniksi

<b>Suomi (konsepti)</b>	<b>Englanti (konsepti)</b>
kokoonpano-ohje	assembly instructions
alumiiniputki	aluminum pipe
pöytäkirja	meeting minutes / minutes of meeting
esijäähdyttää	precool

Kuten kuvasta voi nähdä, lähes kaikki suomenkielessä yhdellä sanalla ilmaistavat konseptit, kirjoitetaan englanninkielessä kahdella tai useammalla sanalla. Tunnisteiden tarkoitus on kuvata tiedostoa mahdollisimman, jonka takia tiedostoa ei voida kuvata aina ainoastaan yhdestä sanasta muodostuvalla tunnisteella. Yksittäisillä sanoilla dokumenttien tunnistaminen on vaihtoehto, joka ei kuitenkaan erota dokumentteja kovin hyvin toisistaan. Sanoja yhdistämällä voidaan muun muassa erottaa alumiiniputket muista putkista, joka on tarpeellista silloin, kun halutaan löytää tietoa ainoastaan alumiiniputkista. Tarkempien hakutulosten saavuttamisen vuoksi Bluefors Oy:n käyttöön oli päätettävä tapa, jolla sanat voidaan yhdistää toisiinsa tunnisteissa.

Guidelines for Tagging-oppaassa tunnisteiden yhdisteleminen toisiinsa mukailee muun muassa sosiaalisessa mediassa nykyään laajalti käytettyjen hashtagitunnisteiden muodostamisperiaatetta. Hashtagit eli, muun muassa Instagramissa käytettävät aihetunnisteet, muodostuvat #-merkistä sekä sanasta, sanoista, numeroista tai muista sallittua merkistöä noudattavasta merkkijonosta, joka on kirjoitettu risuaitamerkin jälkeen ilman välilyöntejä (Finnish Medical Network 2019). Aihetunnisteissa ei sallita välilyöntejä kesken hashtagia, jonka vuoksi sanat on yhdistettävä toisiinsa jollain tavalla. Useasti sosiaalisen median useammasta sanasta koostuvissa aihetunnisteissa, sanat on kirjoitettu yhteen ilman välimerkkejä. Toinen Instagramissa jokseenkin yleinen käytäntö on yhdistää sanat toisiinsa alaviivalla. Halusin valita näistä tavoista yhden, jota noudatetaan yhdenmukaisesti ja sen avulla ehkäistään useampien tunnistevariaatioiden muodostumista.

Vaihtoehdot useampisanaisten tunnisteiden muodostamiseen olivat sanojen yhteen kirjoittaminen sekä alaviivan käyttäminen sanojen välillä. Valitsin näistä kahdesta vaihtoehdosta sen, joka on helpompi lukea. Tämän ajatuksen pohjalta määrittelin tunnisteoppaaseen, että samaan konseptiin kuuluvat sanat yhdistetään toisiinsa alaviivaa käyttäen. Alaviivan käyttö on tyypillistä myös tietotekniikan parissa, jossa sitä käytetään usein yhdistämään monesta sanasta muodostuvan tiedostonnimen osat toisiinsa silloin, kun välilyöntiä ei voida käyttää erottelemaan sanoja (Kotimaisten kielten keskus 2015).

Alaviivalla sanojen yhdistäminen ei päde luonnollisesti niissä tapauksissa, kun englanninkieliset sanat voidaan kirjoittaa yhteen. Alla olevasta taulukosta näkee, miten tunniste-sanat muodostuvat erilaisissa tapauksissa (taulukko 11). Taulukosta voi nähdä myös, että of-sanaa ei toivota käytettävän tunnisteiden muodostamisessa. Oppaaseen määriteltiin nimittäin, että of-sanan sisältämät tunnisteet tulee mahdollisuuksien mukaan muodostaa toisella oikeaoppisella kirjoitustavalla. Muoto, josta puuttuu of-prepositio, on monesti lyhyempi ja jättää vähemmän mahdollisuuksia kirjoitusvirheille.

Taulukko 11: Tunnisteiden muodostaminen alaviivaa käyttäen

Suomi (konsepti)	Englanti (konsepti)	Englanti (tunniste)
kokoonpano-ohje	assembly instructions	Assembly_instructions
alumiiniputki	aluminum pipe	Aluminum_pipe
pöytäkirja	meeting minutes / minutes of meeting	Meeting_minutes
esijäähdyttää	precool	precool

#### 4.4 Tunnistekategoriat

Tunnistekategoriat eivät kuulu Guidelines for Tagging oppaan ensimmäiseen päälukuun, mutta ne ovat olennainen apu tunnisteiden muodostamiseen. Tunnistekategoriat löytyvät oppaan loppuosasta Existing tags -luvusta, johon on kerätty muun muassa määrittelyprosessin aikana sekä kevään 2018 tunnisteprojektin aikana hyödylliseksi todetut tunniste-esimerkit.

Tunnistekategoriat ovat ryhmiä, joiden avulla tunnisteita voidaan lajitella erilaisiin luokkiin. Nämä muistuttavat jonkin verran tietoperustassa esiteltäviä metatietojen lajeja, joilla jokaisella on oma funktio. Tunnisteita on vaikea kategorisoida virallisesti, koska ne voivat olla käytännössä mitä vain tunnisteiden luojasta riippuen (Smith 2007). Tunnistekategorioita voidaan kuitenkin hyödyntää organisaation sisäisessä dokumentaationhallinnassa ohjaamaan, millaisia tunnisteita tiedostoissa toivotaan käytettävän. Samalla nämä kategoriat

toimivat muistutuksena siitä, että tiedostoon on hyvä liittää monipuolisia tunnisteita kuvaamaan dokumentin sisällön lisäksi sen käyttöön oikeuksiin sekä statukseen liittyviä tietoja. Alla on lueteltu tunnisteoppaaseen määritellyt tunnistekategoriat sekä tiivistetty suomeksi niiden kuvaukset (taulukko 12). Nämä kuvaukset nähtävissä englanninkielellä Guidelines for Tagging -oppaasta

Taulukko 12: Tunnistekategoriat ja niiden selitykset

<b>Käyttäjäosasto / Käyttäjäryhmä</b>	Nämä tunnisteet kuvaavat niitä käyttäjäosastoja sekä -ryhmiä, joiden käyttöön tiedosto on luotu. Dokumentilla voi olla useampi tähän kategoriaan kuuluva tunniste, mikäli dokumentti on osoitettu useamman käyttäjäosaston tai -ryhmän käyttöön.
<b>Käyttöoikeudet</b>	Tähän kategoriaan kuuluvat tunnisteet kuvaavat, miten dokumenttia saa julkaista. Dokumentin ollessa salassa pidettävä, siihen tulee liittää tunniste ”salassa pidettävä”.
<b>Status</b>	Status-kategorian tunnisteilla voidaan kuvata dokumentin sen hetkistä tilaa. Tiedosto voi olla esimerkiksi vanhentunut, luonnos tai kopio, joka vaikuttaa dokumentin käyttöön. Vanhentuneita ja luonnosversiossa olevia dokumentteja ei tule julkaista, joka halutaan tuoda muiden käyttäjien tietouteen tämän kategorian tunnisteilla.
<b>Kieli</b>	Kielitunnisteiden avulla käyttäjät voivat merkitä ja löytää dokumentteja, jotka sisältävät tekstiä muulla kuin pelkällä englanninkielellä.
<b>Dokumenttityyppi</b>	Nämä tunnisteet kertovat, millaisesta dokumentista tiedostossa on kyse. Tällaista tietoa voidaan tarvita esimerkiksi silloin, kun käyttäjä haluaa löytää kaikki yrityksen some-materiaalit tai työohjeet yhdellä kertaa.
<b>Aiheet</b>	Aihetunnisteet kuvaavat tiedoston sisältöä ja niitä aiheita, joita dokumentti käsittelee. Tähän kategoriaan voidaan luokitella myös kaikki tunnisteet, jotka eivät sovi muihin tunnistekategorioihin.

Tunnistekategorioiden selitysten lisäksi tunnisteoppaaseen on koottu jokaisen tunnistekategorian alle lista esimerkkitunneista. Näin oppaan lukijan on helppo nähdä nopealla vilkaisulla, mitä ja millaisia tunnisteita kuhunkin kategoriaan kuuluu. Tästä listasta jokainen

voi halutessaan poimia valmiita tunnisteita dokumenttien metatietoihin tai käyttää listaa apuna uusien tunnisteiden luomiseen muut Guidelines for Tagging -oppaan määritelmät mielessä pitäen.

## **5 Kehittäminen ja arviointi**

Tämä luku käsittelee määrittelyprosessin kehittämistä ja arvioimista sekä opinnäytetyön itsearviointia. Luvun ensimmäisen ja toinen aliluku käsittelevät prosessi sekä tunnistepöytä ja voi antaa lukijalle ideoita samankaltaisen toimeksiannon toteuttamiseen

### **5.1 Prosessin kehityskohtia**

Virallisen toimintakielen ja dokumenttienhallintatyökalun määrittelemisen puute toivat prosessiin haastetta, mutta antoivat myös tilaa luovuudelle ja oppimiselle. Kielivalinnan jääminen minulle pakotti minut tutkimaan yrityksessä käytettävää kieltä, jonka aikana tuli esille useita kieleen liittyviä kehityskohtia. Näiden kehityskohtien noustessa esille, esimerkiksi markkinointiosasto alkoi pohtimaan markkinoinnissa käytettävää kieltä ja alkoivat määrittelemään kohdeyrityksen ulkoisessa viestinnässä käytettävää kieltä. Eri kirjoitustapojen esiintyminen toiminnanohjausjärjestelmässä on otettu nyt myös huomioon esimerkiksi uusille työntekijöille suunnatuissa ERP-koulutuksissa. Olen aloittanut myös itse kiinnittämään enemmän huomiota kirjoitustapaan kielen tutkimisen ansiosta. Nykyään olen paljon tietoisempi siitä, milloin käytän amerikanenglantia ja milloin brittienglantia kirjoittaessani englanninkielistä tekstiä.

Työkalun puuttuminen prosessin aikana tarjosi mahdollisuuden tutkia dokumenttienhallintaa laajemmin ja ymmärtää nyt paremmin erilaisten dokumenttienhallintatapojen käyttötarpeet sekä vahvuudet ja heikkoudet. Mikäli työkalu olisi ollut valittu valmiiksi, kaikki määritelmät olisi tehty kyseisen työkalun rajoitteiden puitteissa. Sain siis enemmän vapauksia määritelmien tekemiseen sekä mahdollisuuden miettiä kohdeyrityksen kannalta parhaita ratkaisuja.

Tällaisen prosessin toteuttaminen vaatii toimeksiannon saajalta ehdottomasti hyvää ymmärrystä kohdeyrityksen dokumenteista. Tästä näkökulmasta, oli viisasta, että minä tein kaikki määrittelyt. Minun käyttöoikeudet eivät kuitenkaan riittäneet kaikkien dokumenttien näkemiseen, jonka vuoksi en pystynyt luomaan opasta, jossa kaikkien osastojen dokumentit olisi otettu huomioon. Käyttäjäoikeuksien puolesta projektiin olisi ollut hyvä saada mukaan sellainen henkilö, jolla on oikeudet kaikkien dokumenttien tutkimiseen. Toinen hyvä vaihtoehto olisi ollut perustaa sellainen tiimi määrittelyprosessille, jossa olisi pystytty jakamaan prosessin kannalta tarvittavia tietoja määrittelyprosessia tekevälle henkilölle.



Tällöin Guidelines for Tagging -oppaaseen oltaisiin voitu myös luoda kattavampi listaus olemassa olevista valmiista tunnisteista.

## **5.2 Projektinhallinta**

Opinnäytetyön projektinhallinta ei onnistunut mielestäni niin hyvin kuin olin toivonut. Huomasin etenkin raportin kirjoitusprosessin loppupuolella, että olisin saanut tuotoksen valmiiksi nopeammin, mikäli olisin tutkinut aiheen tietoperustaa tarkemmin jo aloittaessani määrittelyprosessin työstämisen kevään alussa. Työn loppumetreillä hankkimani teoria-tieto metatiedoista aiheutti sen, että Guidelines for Tagging -oppaan sisältöön tuli muutoksia vielä suunnitellun viimeisen version jälkeen.

Minun olisi pitänyt myös suunnitella projektin aikataulu huolellisemmin opinnäytetyön suunnitteluvaiheessa. En pystynyt keskittymään määrittelyprosessiin aina suunnitelmani mukaisesti, koska työskentelin koko prosessin ajan kohdeyrityksessä täysipäiväisesti. Minun täytyi toisinaan priorisoida muita työtehtäviä, niin että prosessi joutui odottamaan. Olin voinut aikatauluttaa prosessin suunnitelmaan enemmän aikaa muiden töiden tekemiselle sekä tiedonhankinnalle, joka oli yllättävän haastavaan tähän aiheeseen liittyen. Olin tehnyt aikataulusta niin tiukan, että ei siihen mahtunut mukaan normaalin elämän käänteitä. Tästä syystä raportin palautus viivästyi kolme kuukautta myöhässä suunnitellusta aikataulusta.

## **5.3 Määrittelyprosessin ja oppaan arviointi**

Määrittelyprosessi oli ehdottomasti haastava minulle, koska minulla ei ollut sosiaalisen median ulkopuolelta minkäänlaista ymmärrystä tunnisteiden käyttämisestä. Aiheen haasteellisuus teki projektista kuitenkin hyvin mielenkiintoisen, sillä sain useita mahdollisuuksia oppia uutta ja kehittää osaamistani minua kiinnostavalla aihealueella. Prosessin aikataulun venyminen johtui pitkälti tiedon etsimiseen kuluneesta ajasta, jota en osannut ottaa huomioon tarpeeksi hyvin opinnäytetyönsuunnitteluvaiheessa. Olin kevään alussa määrittellyt itselleni prosessin valmistumisajaksi elokuun lopun, joka kuitenkin venyi lokakuun loppuun. Projektin viivästyminen ei johtunut ainoastaan haasteista ymmärtää tunnisteisiin liittyviä teknisiä ja kirjoitusasullisia yksityiskohtia, vaan muiden kohdeorganisaatiossa vastuullani olleiden projektien siirtymisestä prioriteeteissa korkeammalle. Kahden kuukauden viivästys ei kuitenkaan viivästyttänyt kokonaisuudessaan dokumenttienhallintatyökalun käyttöönottoa, koska potentiaalisia dokumenttienhallintatyökaluja alettiin tutkia vasta lokakuun puolen välin jälkeen.

Mielestäni oppaan rakenne sekä sen asiasisältö ovat selkeät ja kattavat. Olen sisältöön melko tyytyväinen. Opas on hieman tylsän näköinen ja vaatii mielestäni lisää visuaalisuutta. Oppaan epävisuaalisuus tuli minulle vastaan saadessani palautetta opinnäytetyönohjaajaltani opinnäytetyön suunnitelmasta. En ollut siihen mennessä ajatellut juuri olenkaan, sitä miltä opas näyttää. Oppaan ulkoasu pohjautui silloin ainoastaan Bluefors Oy:n käytössä olevaan työohjedokumenttipohjaan. Paljon tekstiä ja taulukoita sisältävä Guidelines for Tagging -ohje, on edelleen ulkonäöltään melko mitäänsanomaton lukuun ottamatta oppaan viimeisen sivun pikaopasta, josta on liitetty kuva tämän raportin loppuun (Liite 2). Tämä pikaopas lisättiin vasta ymmärrettyäni, miten tylsän näköinen ohjeen työversio oli. Visuaalisia ratkaisuja voisi ehdottomasti kehittää dokumenttiin sen käyttökokeuksen parantamiseksi ja elävöittämään tekstirivejä. Visuaalisuudesta ei tullut mitään kommentteja toimeksiantajalta, joka oli oppaaseen sellaisenaan tyytyväinen viimeisen palautuksen jälkeen.

## Lähteet

Aurinkomatkat. Luettavissa: <http://www.aurinkomatkat.fi/Images/Intian%20viisumiohje.pdf>.  
Luettu 26.10.2019.

Arter Oy. 2019. Opi Lean 5S-projektin suunnittelu, toteutus ja arviointi. Luettavissa: [https://www.arter.fi/qf-koulutukset/opi-lean-5s-projektin-suunnittelu-toteutus-ja-arviointi/?gclid=Cj0KCQjwoqDtBRD-ARIsAL4pviDVm8BGeW1IS0iRTmn6h1YQy-0QAmM-HHfYuVGLs4VZopKf8IBFtjVkaAqM9EALw\\_wcB](https://www.arter.fi/qf-koulutukset/opi-lean-5s-projektin-suunnittelu-toteutus-ja-arviointi/?gclid=Cj0KCQjwoqDtBRD-ARIsAL4pviDVm8BGeW1IS0iRTmn6h1YQy-0QAmM-HHfYuVGLs4VZopKf8IBFtjVkaAqM9EALw_wcB). Luettu: 17.10.2019.

Beare, K. 28.7.2019. ThoughtCo. American English to British English Vocabulary. Luettavissa: <https://www.thoughtco.com/american-english-to-british-english-4010264>. Luettu: 25.10.2019.

Bluefors Oy. 2019a. News. Luettavissa: <https://bluefors.com/opening-a-new-bluefors-office-in-germany/>. Luettu: 5.5.2019.

Bluefors Oy. 2019b. Open positions. Bluefors service hub to US. Luettavissa: <https://bluefors.com/open-positions/#tab-id-17>. Luettu: 24.10.2019.

Bluefors Oy. 2019c. Company. Luettavissa: <https://bluefors.com/company/>. Luettu 5.5.2019.

Bluefors Oy. 2019d. Intranet. Ei luettavissa. Luettu: 16.10.2019.

Bluefors Oy 2019e. Fileserver-tiedostopalvelin. Ei luettavissa. Luettu: 25.4.2019.

Bluefors Oy. 2019f. Lean-toiminnanohjausjärjestelmä. Ei luettavissa. Luettu: 11.4.2019.

British Council. 2019. Differences between British and American English. Luettavissa: <https://www.britishcouncilfoundation.id/en/english/articles/british-and-american-english>.  
Luettu: 25.10.2019

Chui, M. Manyika, J. Bughin, J. Dobbs, R. Roxburgh, C. Sarrazin, H. Sands, G. & Westergren, M. 2012. McKinsey Global Institute. Unlocking value and productivity through social technologies. Luettavissa: <https://www.mckinsey.com/~media/McKinsey/Industries/Tech->

nology%20Media%20and%20Telecommunications/High%20Tech/Our%20Insights/The%20social%20economy/MGI\_The\_social\_economy\_Full\_report.ashx. Luettu 14.11.2019

Cornell University Library. 2002. Research Department. Luettavissa: <http://preservationtutorial.library.cornell.edu/metadata/table5-1.html>. Luettu: 9.10.2019.

CSC – Tieteen tietotekniikan keskus Oy. 2019. Sanasto. Luettavissa: <https://digitalpreservation.fi/specifications/sanasto>. Luettu: 9.10.2019.

Dalkir, K. 2011. Knowledge Management in Theory and Practice. The MIT Press. Lontoo.

Digi-kuva. 2018. Kameramme paljastavat meistä enemmän kuin tiedämme. Luettavissa: <https://digi-kuva.fi/kuvankasittely/metadata-kamera-paljastaa-meista-enemman-kuin-uskoisimme>. Luettu 9.10.2019.

Di Noi, D. 22.5.2018. Blogi. Do workers still waste time searching for information?. Luettavissa: <https://blog.xenit.eu/blog/do-workers-still-waste-time-searching-for-information>. Luettu: 17.10.2019.

Finnish Medical Network. 2019. Aihetunniste eli hashtag (#). Luettavissa: <https://www.fimnet.fi/kaytto-ohjeet/aihetunniste-eli-hashtag/>. Luettu 27.10.2019.

Haastattelut. 02/2019.

Hiltunen, T. 6.8.2019. Välipalaute. Guidelines for Tagging -opas.

IITC. 2019. Dokumenttien hallinta. Luettavissa [https://iitc.fi/dokumenttien\\_hallinta](https://iitc.fi/dokumenttien_hallinta). Luettu 15.6.2019.

Kaario, K. & Peltola, T. 2008. Tiedonhallinta - Avain tietotyö tuottavuuteen. WS Bookwell. Porvoo.

Kankaanpää, L. 2017. Instagram. Luettavissa: [https://www.instagram.com/p/BWTJh6KFDsZ/?utm\\_source=ig\\_web\\_copy\\_link](https://www.instagram.com/p/BWTJh6KFDsZ/?utm_source=ig_web_copy_link). Luettu:20.11.2019.

Keskustelu. 12.6.2019. Bluefors Oy:n toimistoassistentti ja markkinointiasiantuntija.

Keskustelu. 9.9.2019. Bluefors Oy:n työntekijä.

Keskustelu. 16.9.2019. Bluefors Oy:n IT-työntekijä.

Kotimaisten kielten keskus. 2015. Alaviiva. Luettavissa: <http://www.kielitoimistonohjepankki.fi/haku/alaviiva/ohje/7>. Luettu 27.10.2019

Lager, L. 4.5.2016. Metadata ja hyvät käytännöt. OJS-työpaja. Tieteiden talo. Luettavissa: <http://www.kotilava.fi/sites/kotilava.fi/files/2016-5-4-lager.pdf>. Luettu: 9.10.2019.

Lean Lion Oy. 2019. Miksi 5S?. Luettavissa: <https://www.leanlion.com/miksi-5s>. Luettu: 17.10.2019.

Learning English. 2017. Luettavissa: <https://learningenglish.voanews.com/a/six-difference-between-british-and-american-english/3063743.html>. Luettu: 24.10.2019.

M-Files. 2019a. Organize with metadata. Luettavissa: <https://www.m-files.com/en/organize-with-metadata>. Luettu: 19.10.2019.

M-Files. 2019b. Metadata management. Luettavissa: <https://www.m-files.com/fi/metadata-management>. Luettu 26.8.2019.

Microsoft. 2019. View or change the properties for an Office file. Luettavissa: <https://support.office.com/en-us/article/view-or-change-the-properties-for-an-office-file-21d604c2-481e-4379-8e54-1dd4622c6b75>. Luettu 8.10.2019.

MerlinOne. 2019. Our Insights. What are the Different Types of Metadata (and How are They Used)?. Luettavissa: <https://merlinone.com/types-of-metadata/>. Luettu: 9.10.2019.

Oxford Dictionaries. 2019. Haku: Third party. Luettavissa: <https://www.oxfordlearnersdictionaries.com/definition/english/third-party?q=third+party>. Luettu: 27.10.2019.

Oxford International Education Group. 2019. Luettavissa: <https://www.oxfordinternationalenglish.com/differences-in-british-and-american-spelling/>. Luettu: 24.10.2019

Puura, H. 25.6.2019. Head of Human Operations. Bluefors Oy. Sähköposti.

- Sanastokeskus TSK. 2010. Hakusana: tag. Luettavissa: <http://www.tsk.fi/tepa/fi/haku/tag>.  
Luettu: 12.10.2019.
- Sedu. 2017. Instagramin käytön perusteet. Luettavissa: <https://www.sedu.fi/loader.aspx?id=7f186311-19c9-464d-a594-a26cc06ca837>. Luettu 20.11.2019.
- Serola S. 2010. Ote informaatiosta – Johdatus informaatiotutkimukseen ja interaktiiviseen mediaan. BTJ Kustannus. Helsinki.
- Smith, G. 2007. Tagging: People-powered Metadata for the Social Web. New Riders.  
<https://www.oreilly.com/library/view/tagging-people-powered-metadata/9780321550149/>.  
Luettu: 19.9.2019.
- Smith, G. 2008. Tagging: People-powered Metadata for the Social Web. New Riders.  
Berkley.
- Ståhle, P. & Grönroos, M. 2002. Knowledge management – tietopääoma yrityksen kilpailutekijänä. WS Bookwell Oy. Porvoo.
- Tieteen termipankki. 2016a. Informaation. Luettavissa: <http://tieteentermipankki.fi/wiki/Filosofia:informaatio>. Luettu 14.11.2019
- Tieteen termipankki. 2016b. Metatieto. Luettavissa: <http://tieteentermipankki.fi/wiki/Nimitys:metatieto>. Luettu: 9.10.2019.
- Trends Magazine. 2013. A New Era for "Interaction Workers". Luettavissa: <https://audiotech.com/trends-magazine/a-new-era-for-interaction-workers/>. Luettu 14.10.2019.
- UC Merced Library. 2019. File and folder Organization – long draft. Luettavissa: <http://library.ucmerced.edu/node/66751>. Luettu: 15.6.2019.
- Vähäsarja, S. 27.2.2015. Yle. Kotimaa. Ääkköset nimessä voivat vaikeuttaa matkustamista – yksiselitteistä ohjetta niiden käsittelyyn ei löydy. Luettavissa: <https://yle.fi/uutiset/3-7734064>. Luettu: 27.10.2019.
- Yle. 15.11.2016. Huipputeknologia ja Lapin matkailu palkittiin kansainvälistymispalkinnolla. Luettavissa: <https://yle.fi/uutiset/3-9294459>. Luettu: 16.10.2019.

### Table of Contents

<b>Tags in Bluefors document management.....</b>	<b>1</b>
<b>What is a tag? .....</b>	<b>1</b>
<b>How tags work? .....</b>	<b>1</b>
<b>1 Tag structure and definitions .....</b>	<b>2</b>
<b>1.1 Language: American English.....</b>	<b>2</b>
Non-English and multilingual documents.....	2
<b>1.2 Characters .....</b>	<b>3</b>
1.2.1 Characters to use.....	3
Alphabetic characters .....	3
Letter replacement .....	3
Numeric characters .....	3
Symbols.....	4
1.2.2 Characters not to use .....	4
Symbols.....	4
Ampersand '&' .....	4
<b>1.3 Spelling.....</b>	<b>5</b>
1.3.1 Singular or plural? .....	5
1.3.2 Piling words and characters into a single tag .....	6
1.3.2.1 Why it is usually better not pile multiple words into a tag? .....	6
1.3.2.2 How to combine multiple words into one tag.....	6
<b>2 How to choose correct tags to apply to a document?.....</b>	<b>7</b>
<b>2.1 Key points to keep in mind when creating tags .....</b>	<b>7</b>
<b>2.2 Example cases .....</b>	<b>7</b>
<b>3 Existing tags .....</b>	<b>8</b>
<b>3.1 Tag categories.....</b>	<b>8</b>
<b>3.2 Existing tags by category.....</b>	<b>9</b>
Document type.....	9
Topics .....	10
Status.....	11
Language .....	11
Access rights.....	11
User department / User group.....	12
<b>4 Good tagging process in a nutshell .....</b>	<b>13</b>

